JCTC Journal of Chemical Theory and Computation

TIP5P-Consistent Treatment of Electrostatics for Biomolecular Simulations

Sarah M. Tschampel, Michael R. Kennerty, and Robert J. Woods*

Complex Carbohydrate Research Center, 315 Riverbend Road, Athens, Georgia 30602

Received February 23, 2007

Abstract: The inclusion of zero-mass point charges around electronegative atoms, such as oxygen, within molecular mechanical force fields is known to improve hydrogen-bonding directionality. In parallel, inclusion of lone-pairs (LPs) in the TIP5P water model increased its ability to reproduce both gas-phase and condensed-phase properties over its non-LP predecessor, TIP3P. Currently, most biomolecular parameter sets compute partial atomic charges via fitting of the classical molecular electrostatic potential (MEP) to the quantum mechanical MEP. Application of this methodology to optimize lone-pair description is therefore consistent with the current approach to modeling electrostatics and is straightforward to implement. Here, we present an atom-type specific lone-pair model, which leads to the most optimal LP placement for each atom type, and, notably, results in reproduction of the lone-pair description present in TIP5P. Carbohydrates are rich in hydroxyl groups, and development of a lone-pair inclusive carbohydrate force field for use with a lone-pair containing water model, such as TIP5P, ensures the compatibility between these two models. Implementation of this lone-pair model improves the geometry and energetics for a series of hydrogen-bonded clusters and the properties of several small molecule crystals over the non-LP containing force field.

Introduction

Partial charges are nonphysical entities that are nevertheless convenient to employ in the computation of the nonbonded Coulomb interaction in molecular mechanics force fields, eq 1.[1] The common use of atom-centered partial charges (monopoles) exclusively is based on the approximation that the higher order terms (dipole, quadrupole, etc.) can be ignored, due to the rapid rate at which the higher order contributions diminish with respect to internuclear distance. A variety of protocols have been developed to obtain atom-centered partial atomic charges for biomolecules. Several groups have pioneered the general method that employs the computed quantum mechanical molecular electrostatic potential (MEP) at a grid or shell of points around a given molecule to derive partial charges, eq 2.[1−10] Partial charges are fit to the atomic centers in a given molecule, so as to optimize the agreement between the classical MEP arising from these partial charges and the quantum mechanical MEP, eq 3.[1,7] Least-squares fitting yields a minimized error (χ), which is utilized here to gauge the quality of the fit between the classical and quantum mechanical MEP.

$$V_{\text{electrostatic}} = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{1}$$

$$\text{MEP}_{\text{QM}}(r) = \sum_{A} \frac{Z_A}{r - R_A} - \sum_{\mu,\nu} P_{\mu,\nu} \int \frac{\varphi_\mu \varphi_\nu}{|r - r'|} \, \mathrm{d}r' \tag{2}$$

$$\text{MEP}_{\text{classical}}(i) = \sum_{j=1}^{n} \frac{q_j}{r_{ij}} \tag{3}$$

Despite subtle differences in the derivation of partial atomic charges, most of the current MEP-based techniques have been shown to perform well in practice and are widely employed. In contrast, the inclusion of point charges around oxygen, nitrogen, and sulfur, to mimic lone-pair electrons, continues to be controversial. The option of including lone-

* Corresponding author phone: (706)542-4454; fax: (706)542-4412; e-mail: rwoods@ccrc.uga.edu.

10.1021/ct700046j CCC: $37.00
Published on Web 07/12/2007

pairs in molecular dynamics (MD) simulations with the AMBER program has been available since 1984.[9,11−13] Since 2000, the AMBER package has employed an analytical treatment of lone-pairs with zero-mass and rigid relative positions.[14] Previously, lone-pairs were treated as pseudoatoms with both a mass, a partial charge, and associated valence force constants.[14] Along with the conversion to an analytical treatment of lone-pairs ("Extended Points" (EPs) in AMBER nomenclature) the distance of the lone-pairs from the corresponding oxygen nucleus was only marginally altered from 0.20 Å (PARM81) to 0.35 Å (PARM02EP).[12,15−17] The method used to determine the lone-pair oxygen bond length in PARM02EP was based on the location of critical points in the charge density.[18−20] Critical points were identified for a test set of 21 compounds comprised of sp, $sp,^2$ and sp^3 nitrogen as well as sp^2 and sp^3 oxygen atom types. Due to the insensitivity of critical point position to chemical environment, it was inferred that an average value of 0.35 Å for the LP−O bond length was transferable to all relevant oxygen and nitrogen atom types. More recently, a similar method was utilized for preliminary lone-pair placement to develop a polarizable lone-pair inclusive model.[21] However, the critical point location in the electron density is dependent on the level of theory and basis set employed in the quantum mechanical calculation.[9] Furthermore, the charge density has not been shown to lead to good reproduction of the MEP. Therefore, since fitting partial charges to the quantum mechanical MEP has been shown to be appropriate for condensed-phase MD simulations, a LP placement derived from this approach was examined for application within the AMBER force field for the AMBER and GLYCAM parameter sets.

Empirically adjusting partial charges to reproduce experimentally known condensed-phase properties, such as density and enthalpy of vaporization, from liquid-state MD simulations is an alternative method often used to compute partial charges for pure liquids, such as water, giving rise to the TIP5P water model.[22−25] In the case of TIP5P, the LP placement is specific for the hydroxyl-type oxygen present in the water molecule. This technique was also applied to the derivation of the LP−O distance for a model compound, methanol, in which the oxygen was approximated to be representative for all oxygen atom types. A distance of 0.47 Å was determined and applied to all oxygen atom types in a polarizable version of the OPLS force field.[26] This empirical treatment cannot be applied to amino acids, nucleic acids, or carbohydrates since they do not exist as pure liquids, and it is unclear to what extent values determined for model liquids are transferable to solutes.

Due to the number of hydroxyl groups as well as the presence of a ring oxygen atom carbohydrates inherently possess a large number of potential LP sites; in a typical hexopyranoside, such as methyl α-D-glucopyranoside (α-D-Glc*p*-OMe), there are 12 possible LP placements. Therefore, effects observed for the addition of LPs to one electronegative atom may be amplified in carbohydrates. Due to the number of hydroxyl groups, carbohydrates have the ability to form multiple inter- and intraresidue as well as solute−solvent hydrogen bonds. As a result, inclusion of LPs may have a beneficial effect on hydrogen-bond directionality in carbohydrate-containing systems, such as glycoproteins and carbohydrate−protein complexes. In these systems, the orientation of the glycan relative to the protein surface is directly influenced by intermolecular hydrogen bonds and may further include interactions with bridging water molecules.[27,28]

Derivation of an approach that can predict the LP description in TIP5P, and yet is applicable to biomolecular solutes, would result in a TIP5P-consistent approach to developing a LP-containing biomolecular force field. Currently, the GLYCAM parameter set is consistent with the AMBER methodology for partial charge derivation, which involves determination of partial charges that produce a classical MEP with the best fit to the quantum mechanical HF/6-31G(d) MEP.[16,29−32] Extension of this technique to determine the LP−O distance as well as the LP partial charges will ensure consistency with the current methodology, as long as the fit is to the quantum mechanical MEP computed with the same HF/6-31G(d) wave function. Nevertheless, the effect of the level of theory and basis set utilized to obtain the quantum mechanical MEP on the determination of the LP−O separation will be examined. In contrast to the results from critical point analysis of the charge density, analysis of the fits to the MEPs for a variety of test compounds, which include sp^2 and sp^3 oxygen atoms, show a correlation between the LP descriptors (LP−O distance and partial charge) and chemical environment. Therefore, different oxygen atom types can have different LP descriptors. Lone-pair placement around nitrogen atoms can also be derived with this approach, but addition of lone-pairs attached to oxygen atoms has previously been shown to have a more profound effect on the ability to reproduce characteristics of hydrogen bonding interactions and will only be examined here.[26,33]

Once implemented in the AMBER/GLYCAM parameter set, an analysis of the ability of this model to reproduce experimental and theoretical properties for gas-phase hydrogen-bonded clusters as well as monosaccharides in their crystalline form will be presented.[34,35]

Methods

Quantum Mechanical Computations. All quantum mechanical computations were performed with the Gaussian 98 suite of programs, version A.11.3.[36] The optimized geometry of water was computed at the HF/6-31G(d) and B3LYP/6-31G(d) levels.[37−40] The quantum mechanical MEPs were computed with the grid-based CHarges from the ELectrostatic Potentials (ChelpG) algorithm as implemented in Gaussian 98 at the HF and B3LYP level with the 6-31G(d) basis set.[7] Partial charges were computed utilizing the Restrained Electrostatic Potential (RESP[32]) scheme implementing a weighting factor of 0.010.[35]

A series of model compounds representing the oxygen atom types found in alcohols, ethers, ketones, amides, and carboxylate compounds was optimized, and the quantum mechanical MEPs were determined at the HF/6-31G(d), B3LYP/aug-cc-pVTZ, and MP2/aug-cc-pVTZ levels.[41,42] The use of this basis set with the B3LYP functional is to maintain

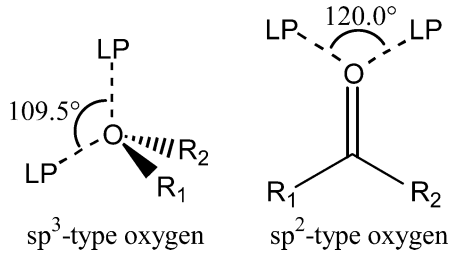

Figure 1. The lone-pair geometry for sp^3 and sp^2 oxygen atoms.

consistency with previously developed models implemented in AMBER.[43] For all sp^3-type oxygen atoms the LP geometry was constrained to prefer a tetrahedral placement, while for all sp^2-type oxygen atoms the LPs were constrained to be in the plane of the carbonyl group and its two substituents, forming a 120° angle, except for the carboxylate compounds, Figure 1.[15,26]

Molecular Mechanics. The GLYCAM06 parameters were utilized in conjunction with the PARM94 parameter set, with the addition of LPs that are defined and fixed in relation to their attached oxygen atom type and have no van der Waals radii. Energy minimizations were performed with the SANDER module of AMBER7 with a dielectric constant of unity for 10 000 cycles (9000 steepest descent followed by 1000 conjugate gradient).[14] All atoms were included in the calculation of nonbonded interactions, and each energy minimization was initiated from the ab initio geometry to which it was compared. Two types of systems were investigated, in which the first type of system consisted of a series of small hydrogen-bonded clusters that did not employ any geometrical restraints and the second investigated the approach of a water molecule toward either methanol or *N*-methyl acetamide (NMA) in which additional restraints were necessary.[44,45] In each of the approach trajectories, the intermolecular O···O distance was restrained at the QM value along the potential energy surface. In addition, the methyl group was restrained from rotating away from the ab initio value in the NMA−water complex with a harmonic restraint. The only additional restraints were employed as needed to ensure the water molecule remained in the QM orientation relative to the methanol or NMA molecule.

Crystal Simulations. The unit cells for the orthorhombic crystals of α-D-Glc*p* (GLUCSA10), α-D-Glc*p*-OMe (MGLUCP11), α-D-Man*p*-OMe (MEMANP), and *β*-D-Gal*p*-OMe (MBDGAL02) were each transformed using the $P2_12_12_1$ symmetry operators to generate crystal lattices containing 64 monosaccharides.[46−48] Similarly, the $P2_1$ symmetry operators were utilized to generate a 2 × 4 × 2 system for α-D-Glc*p* monohydrate (GLUCMH11) and a 4 × 4 × 4 system for α-D-Glc*p*NAc (ACGLUA11).[49,50] The experimental monosaccharide conformation was employed as the initial configuration. Subsequently, no position or symmetry restraints were applied, and the box dimensions were allowed to vary over the course of the MD simulations. Each system governed by the $P2_12_12_1$ symmetry operator was obtained via neutron diffraction allowing for direct determination of the proton positions, while the protons in the $P2_1$ structures were obtained by experimental crystal density difference experiments.

Table 1. Partial Atomic Charges (au) for α-D-Glc*p*

GLYCAM	2000a	06	06-LP
C1	0.151	0.509	0.292
C2	0.131	0.246	0.170
C3	0.211	0.286	0.109
C4	0.160	0.255	0.146
C5	0.085	0.283	0.227
C6	0.244	0.277	0.138
O1/LP	−0.612	−0.639	0.000/−0.220[a]
O2/LP	−0.632	−0.713	0.000/−0.218
O3/LP	−0.668	−0.699	0.000/−0.215
O4/LP	−0.665	−0.710	0.000/−0.214
O5/LP	−0.404	−0.574	0.000/−0.183
O6/LP	−0.671	−0.682	0.000/−0.210
H1	0.153	0.000	0.000
H2	0.103	0.000	0.000
H3	0.061	0.000	0.000
H4	0.081	0.000	0.000
H5	0.086	0.000	0.000
H6/H6′	0.031/0.031	0.000/0.000	0.000/0.000
HO1	0.432	0.445	0.327
HO2	0.415	0.437	0.281
HO3	0.430	0.427	0.280
HO4	0.429	0.436	0.278
HO6	0.417	0.418	0.272

[a] The partial charge on a single LP is listed; note there are 2 LPs per oxygen atom.

For each monosaccharide, ensemble-averaged charges were derived utilizing the same 100 conformers employed previously in the generation of ensemble averaged charges in the GLYCAM2000a parameter set.[31,51] Unlike GLYCAM-2000a, in GLYCAM06 aliphatic hydrogen atoms are excluded from the charge fitting.[30] Exclusion of the aliphatic hydrogen atoms during the charge fitting led to a robust partial charge set, with nearly identical values for the same atom types within the monosaccharide, Table 1.[52] A similar charge fitting scheme was utilized to determine corresponding partial charges for the ala_2 zwitterion.[53] The unit cell of the alanine dipeptide zwitterion contains 8 molecules, as determined by experimental crystal density difference diffraction, and was transformed to a 128 molecule cell utilizing the *I*4 symmetry operators (2 × 2 × 4). The experimental conformation was utilized in the initial configuration, which contained all hydrogen atoms except one methyl hydrogen atom. The missing methyl hydrogen atom was added based on the standard tetrahedral configuration of methyl groups.

Unit scaling factors for 1,4 nonbonded interactions were employed during the MD crystal simulations of carbohydrates, which is consistent with the GLYCAM2000a and GLYCAM06 parameter sets.[30,31] The Particle Mesh Ewald algorithm was implemented for treatment of long-range interactions as each system was heated from 5 to 300 K (experimental temperature range was 283−303 K) over 50 ps and subsequently maintained at 300 K for 1 ns utilizing the Berendsen temperature coupling scheme.[54,55] The effect of both isotropic and anisotropic pressure scaling was implemented to establish the effects these pressure models had on the lattice structure. The deviation from experiment was amplified with the anisotropic model, but the same

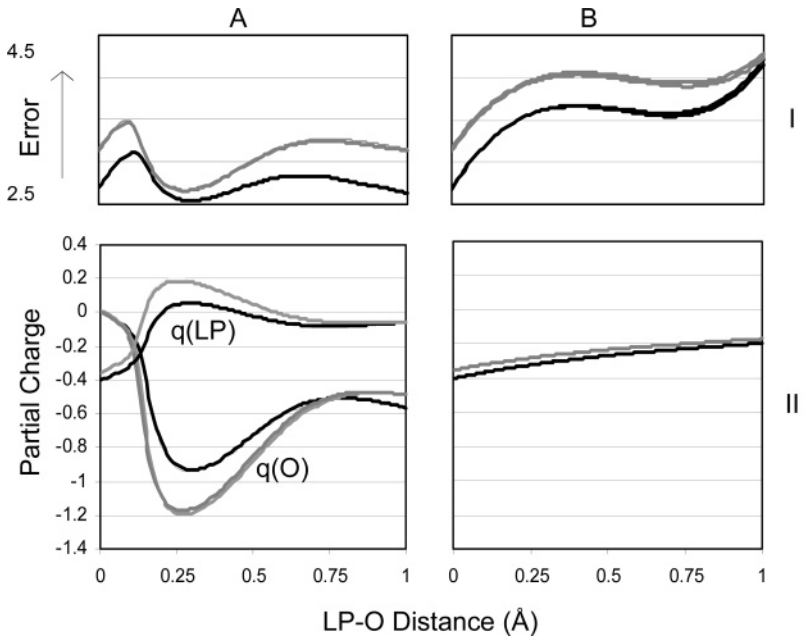

Figure 2. Correlation of the error function (I) and the partial charges (II) in scheme **A** (q(O) and q(LP)) and in scheme **B** (q(LP), q(O) = 0) from fitting to the B3LYP/6-31G (gray) and HF/6-31G(d) (black) MEPs for a water molecule.

relative trends in cell distortion were observed with both models. Due to program limitations, anisotropic scaling was not implemented in the crystals governed by $P2_1$ symmetry. Data from only the last 500 ps were utilized to determine the average cell dimensions and hydrogen-bond distances.[34,35,48]

Results and Discussion

Determination of the Optimal LP−O Distance. Initially, the effect of fitting to the electrostatic potential to determine the placement of LPs around oxygen was assessed by determining the quality of the fit between the classical and quantum mechanical MEPs in a water molecule as a function of lone-pair position. The two LP−O distances were scanned from 0.01 to 1.00 Å, in 0.01 Å increments, with both LP−O distances and the partial atomic charge on the lone-pairs constrained to be equal. Two different approaches were taken during the RESP charge fitting stage. In the first scheme (**A**) a charge on oxygen was permitted, while in the second scheme (**B**) the charge on oxygen was constrained to zero, as in the TIP5P water model. Scheme **A** resulted in a minimum at a short LP−O distance of approximately 0.3 Å but possessed unintuitive partial charges; the LP charge became increasingly positive as the error function was minimized. Application of scheme **B** led to a minimum at a longer LP−O distance of approximately 0.75 Å, with negatively charged lone-pairs of approximately −0.2 *e*, Figure 2. Notably, the LP description obtained from scheme B is very similar to that in the TIP5P water model, which has an LP−O distance of 0.70 Å, and q(O) = 0 *e*, and q(LP) = −0.24 *e*. Thus, the well-defined MEP fitting approach reproduces the electrostatic description of the TIP5P water model, despite the fact that the TIP5P water model was originally obtained via empirical fitting to bulk liquid properties.

Schemes **A** and **B** were applied to a series of test molecules, including alcohols and ethers, to determine the optimal LP−O descriptors for sp^3 type oxygen atoms, Table 2. The average LP−O distance at the best fit was 0.69 Å from the HF/6-31G(d) MEP, 0.72 Å from the B3LYP/aug-cc-pVTZ MEP, and 0.71 Å from the MP2/aug-cc-pVTZ MEP with scheme **A**. Similarly, application of scheme **B** led to average optimal LP−O distances of 0.71, 0.75, and 0.71 Å, respectively. Notably in scheme **A** only a small partial charge was observed on the oxygen atom in the alcohols and ethers. Further, in some cases this small charge was positive. Each of the LP−O distances derived thus far was based on the approximation that the LP−O distances are equivalent, which is consistent with previous studies and molecular symmetry.[15,26,33] When both LP−O distances were varied independently for each of the alcohol and ether model compounds, the optimal locations converged to the symmetric results, Figure 3. Therefore, the results for the sp^3 type oxygen atoms in water, alcohols, and ethers show that the most consistent and intuitive fit between the classical and quantum mechanical MEP was obtained when the partial charge on oxygen was constrained to zero (scheme **B**) and the LPs should have equivalent negative charges, which are determined directly from the RESP-fitting.

In order to extend this model to sp^2-type oxygen atoms, a test set consisting of ketones and amides was compiled. For the ketones, the fit resulted in a very flat error function surface, and, at longer LP−O distances, the fit deteriorated and the partial charges at the lone-pair sites became positive. An optimal fit was maintained at shorter LP−O distances, of less than 0.5 Å, with a shallow minimum at approximately 0.3 Å. In addition, at shorter LP−O distances the charge on oxygen was nearly zero, even though it was allowed to vary, and the lone-pairs maintained a negative charge in scheme **A**. Therefore, the optimal lone-pair placement for ketones

Table 2. LP−O Distance, *q*(O), and *q*(LP) at the Best Fit between the Classical and ab Initio MEP for Each Molecule in the sp^3 Oxygen Atom Test Set at the HF/6-31G(d) (I), B3LYP/aug-cc-pVTZ (II), and MP2/aug-cc-pVTZ (III) Levels

	q(O) ≠ 0 (scheme **A**)			*q*(O) = 0 (scheme **B**)	
	LP−O	*q*(O)	*q*(LP)	LP−O	*q*(LP)
methanol					
I	*0.70*[a]	−0.029	−0.178	0.80	−0.177
II	*0.70*	0.017	−0.169	0.88	−0.149
III	0.89	−0.046	−0.146	0.84	−0.163
ethanol					
I	0.65	−0.047	−0.195	0.57	−0.221
II	0.74	−0.043	−0.168	0.68	−0.188
III	0.71	−0.048	−0.179	0.64	−0.202
2-butanol					
I	0.53	0.008	−0.240	0.53	−0.237
II	0.58	−0.005	−0.205	0.56	−0.209
III	0.53	0.001	−0.228	0.52	−0.229
2-propanol					
I	*0.70*	−0.097	−0.169	0.70	−0.199
II	*0.70*	−0.074	−0.157	0.77	−0.173
III	*0.70*	−0.094	−0.163	0.72	−0.190
3-pentanol					
I	0.59	0.031	−0.232	0.61	−0.218
II	0.65	0.013	−0.197	0.66	−0.191
III	0.60	0.014	−0.216	0.60	−0.213
t-butanol					
I	*0.70*	−0.080	−0.173	0.76	−0.192
II	*0.70*	−0.076	−0.160	0.80	−0.174
III	*0.70*	−0.093	−0.170	0.77	−0.192
dimethyl ether					
I	0.92	−0.011	−0.105	0.91	−0.115
II	0.90	0.110	−0.129	0.70	−0.106
III	0.81	0.183	−0.168	0.70	−0.116
ethyl methyl ether					
I	0.70	0.052	−0.150	0.80	−0.126
II	0.79	0.158	−0.177	0.88	−0.109
III	0.74	0.188	−0.188	0.91	−0.114
methyl isopropyl ether					
I	0.65	0.047	−0.163	0.67	−0.145
II	0.72	0.097	−0.167	0.79	−0.131
diisopropyl ether					
I	0.73	0.041	−0.160	0.74	−0.147
II	0.75	0.049	−0.158	0.76	−0.143

[a] Italics indicate systems in which the resolution was not fine enough to detect a precise minimum, so the partial charges at an LP−O distance of 0.70 Å are shown.

was found to be 0.3 Å from the oxygen atom and the partial charge on the oxygen set to zero.

In contrast to all the model compounds investigated thus far, inclusion of symmetric lone-pairs around the carbonyl oxygen of the amide group slightly deteriorated the fit to the MEP containing molecules. Allowing each lone-pair to adopt a unique distance from the carbonyl oxygen atom improved the fit for the amides, with the best fit at a combination of a longer (0.7 Å) LP1−O separation and a shorter (0.3 Å) LP2−O separation, Figure 4. The shorter LP2−O distance is located on the side of the carbonyl containing the amide nitrogen.

The charge model for anionic carboxylate groups, such as those in aspartate and glutamate, did not incorporate lone-pairs in the PARM02EP parameter set.[15] In order to examine the validity of this approximation several different lone-pair placements were investigated for acetate, propanoic acid anion, 2-methylbutanoic acid anion, and 2-methylpropanoic anion. Initially, the carboxylate group was treated as containing two ketone oygen atoms, and the standard sp^2 geometrical placement was applied, the LP−O−LP angle of 120° as in Figure 5 (I). The two lone-pairs located between the oxygen atoms in the carboxylate group (LP_{inner}) and the remaining two lone-pairs (LP_{outer}) were constrained to have equivalent partial charges and separation from their respective oxygen atoms to maintain molecular symmetry. Second, the two LP_{inner} partial charge sites were replaced by a single partial charge, Figure 6(II).

Third, removal of both LP_{outer} sites from the second model, leaving only a single LP bisecting the O−C−O angle was investigated, Figure 5(III), which would theoretically help to alleviate the imbalance observed during MD simulations. Commonly, a hydrogen bond formed with a carboxylate entity is bifurcated in nature, but during an MD simulation at 300 K the hydrogen bond donor typically associates with a single oxygen atom at a given point in time. Overall, the majority of the negative charge is located on the two oxygen atoms, while the charge on LP_{inner} is only slightly negative, ranging only to −0.03 *e* for the HF and B3LYP levels, and becomes positive for all LP_{inner}−C distances less than 2 Å at the MP2 level for this model, Figures 5 and 6(III). Since the inclusion of LPs in anionic carboxylate groups did not lead to an improvement in the fit between MEPs, nor to a minimum in the error function, lone-pairs were not included in carboxylate groups.

These atom-type specific lone-pair placements for oxygen will be applied to GLYCAM06, and ensemble averaged partial charges will be derived for each pyranoside to yield GLYCAM06-LP.

Analysis of Hydrogen-Bonded Clusters. In order to assess the accuracy of the GLYCAM06-LP model, the geometry of several hydrogen-bonded clusters were examined in vacuo. Each cluster contained at least one water molecule, which, if no LPs were present, was modeled as TIP3P or, when the new LP model was employed, as TIP5P. The B3LYP/6-31+G(d) optimized geometry was used as the starting point for each cluster geometry force field energy minimization.[45] The root-mean-square deviations (RMSDs) for both the heavy atoms and for all atoms in each neutral cluster were determined between the energy minimized molecular mechanics structure and the quantum mechanically optimized geometry. Overall, for the 21 clusters examined, the RMSDs for the non-LP containing model and the new LP model were 0.58 Å (0.30 Å) and 0.49 Å (0.21 Å), respectively (heavy-atom RMSDs shown in parentheses).

Examination of the relative energies for the approach of water to a hydrogen acceptor or donor-containing molecule provides a useful test of the electrostatic model and illustrates the applicability of the chosen water model. For the methanol−water cluster, the approach in which methanol is the hydrogen bond donor (MdW) and the approach in which

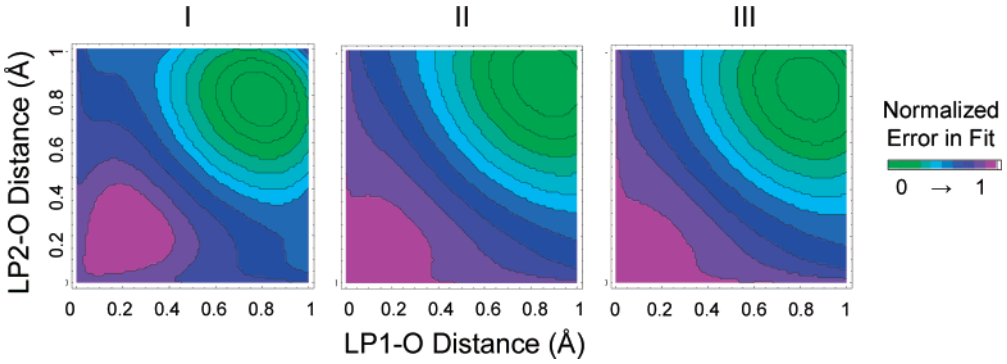

Figure 3. Variation of χ with respect to the LP−O distance at the HF/6-31G(d) (I), B3LYP/aug-cc-pVTZ (II), and MP2/aug-cc-pVTZ (III) levels with the partial charge on oxygen set to zero for methanol.

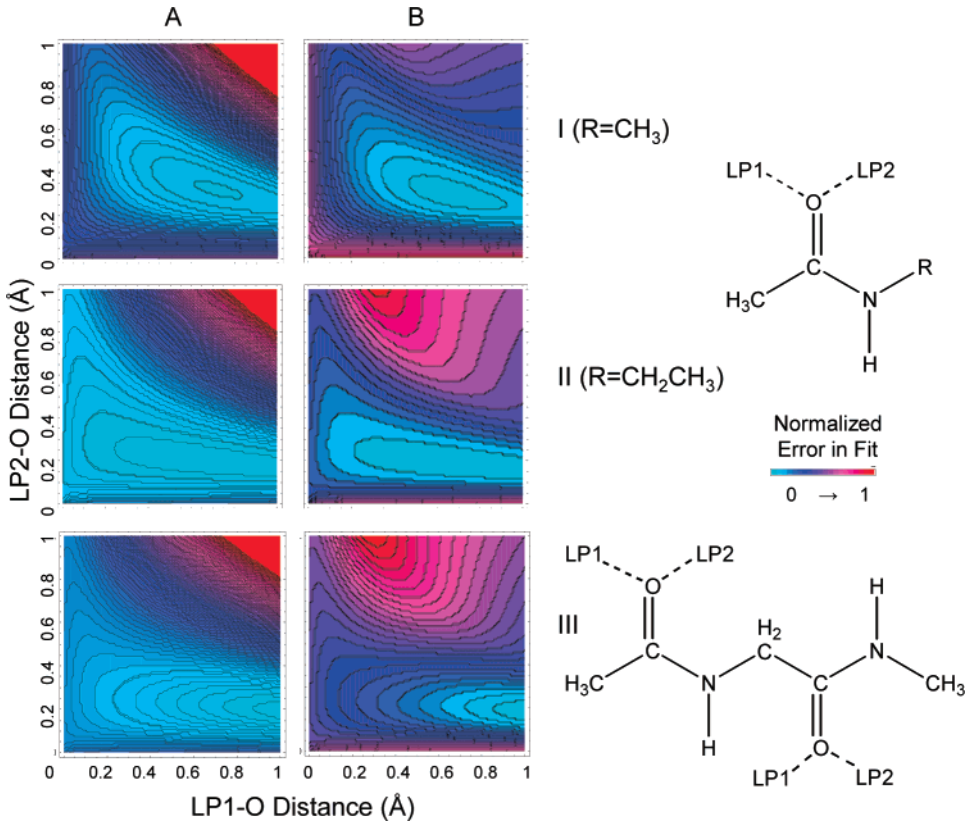

Figure 4. Variation of χ with respect to the LP−O distance (Å) in *N*-methyl acetamide (I), and *N*-ethyl acetamide (II), and the glycine dipeptide (III). The partial charge on oxygen was freely determined in scheme **A**, while the oxygen partial charge was set to zero during the RESP fit in scheme **B**.

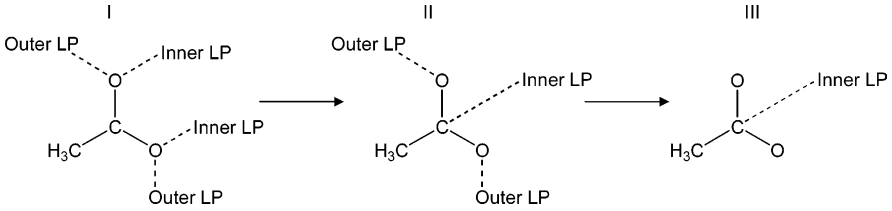

Figure 5. LP placement for anionic carboxylate groups.

water is the hydrogen bond donor (WdM) were compared at the HF/6-31G(d) level, which is the same level at which the MEPs were determined, Figure 7. At this level of theory there is no difference for the two configurations.[44] At the global minima, determined at the MP2/aug-cc-pVQZ level, the interaction energy for the WdM dimer (−5.72 kcal·mol^{-1}) was more favorable than for MdW (−4.95 kcal·mol^{-1}).[44] Although both the non-LP and LP models reversed the relative ranking of the two configurations, the approach trajectory of the LP model was energetically much closer than the non-LP model. In both cases the energy difference was less than 1 kcal·mol^{-1}, specifically, 0.6 and 0.3 kcal·mol^{-1} for the non-LP and LP models, respectively. Further work, such as the inclusion of a van der Waals term

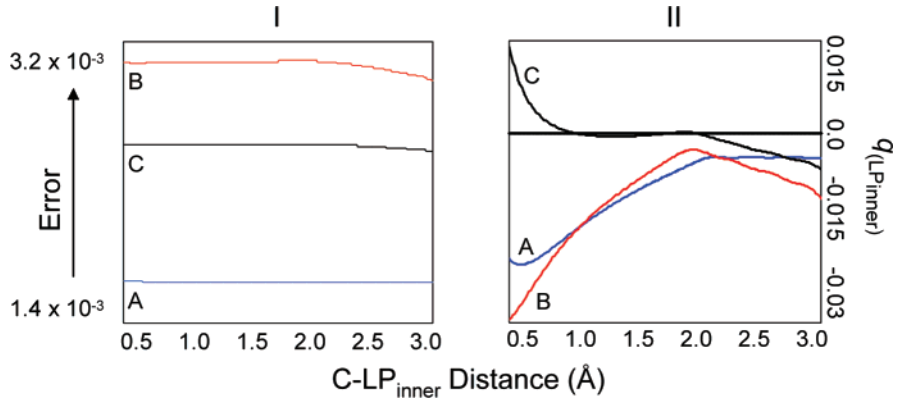

Figure 6. Variation of χ (I) and charge on LP_{inner} (II) with respect to the $LP_{inner}-C$ distance when the QM level was HF/6-31G(d) (A, blue), B3LYP/aug-cc-PVTZ (B, red), or MP2/aug-cc-PVTZ (C, black) for acetate for charge arrangement III in Figure 5.

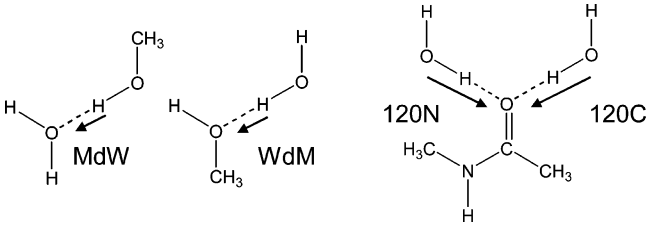

Figure 7. Approach of H_2O to the symmetrical sp^3-type lone-pair hydroxyl group in methanol (left) and the approach of H_2O from the each of the asymmetric sp^2-type lone-pair axes of NMA (right).

on the hydroxyl hydrogen may be necessary to reproduce the high-level quantum results. For example, inclusion of a van der Waals parameter on the hydrogen atoms increased the ability of the TIP3P water model to reproduce bulk water properties when the particle mesh ewald algorithm was employed.[56]

The comparison of the relative energies associated with the approach of a water molecule along each of the LP−O axes in *N*-methyl acetamide ensures that the asymmetric lone-pair model degrades neither the geometry nor the energetics of the hydrogen-bonded cluster. Approach along the LP−O axes from the methyl side is more favorable than the approach from the *N*-methyl side (120N), by 1.3 kcal·mol^{-1} at the B3LYP/6-31+G(d, p) level. Starting from the quantum mechanically determined structures, energy minimization utilizing the non-LP containing model with TIP3P reverses the ranking obtained with density functional theory, while inclusion of LPs and the TIP5P water model reproduces the quantum mechanically determined ranking, Figure 7.

MD Crystal Simulations. MD simulations of monosaccharide crystals provide a sensitive method for testing nonbonded parameters in the condensed phase and have been performed previously for carbohydrates.[34,35] If the nonbonded parameters are too attractive, then it leads to a more tightly packed crystal, while if the interactions are underestimated, then the crystal cell will expand. The addition of lone-pairs is purely a variation in the electrostatic model, and monitoring the behavior of cell lengths over the course of MD simulations, for both the non-LP and LP models, provides a sound test for the effects of lone-pair inclusion.

Initially, for each of three methyl glycosides, α-D-Glc*p*-OMe, *β*-D-Gal*p*-OMe, and α-D-Man*p*-OMe, a 64 and a 256 molecular unit system was generated. Two different size systems were employed with the GLYCAM06 parameter set to determine if the smaller system size introduced any artifacts in the distortion of the cell dimensions. The average difference in cell lengths for the three methyl glycosides in the large (256 molecules) cell relative to the 64 molecule cell was 0.01%. Nonbonded cutoffs of 8, 9, and 10 Å were investigated within the 256 molecule cells, yielding less than a 0.02 and 0.03% difference for the 9 and 10 Å cutoff, respectively, relative to the 8 Å cutoff for the three methyl glycosides. A time step of 0.5 fs was implemented as a standard to determine if longer time steps would be appropriate. Due to the negligible difference in the results obtained with a time step of 1 fs relative to the 0.5 fs initial time step, the larger time step was employed along with an 8 Å cutoff and the smaller lattice size.

The majority of the improvement observed from utilization of the GLYCAM2000a to the GLYCAM06 and 06-LP force fields can be accounted for mainly by the change in nonbonded parameters for the hydroxyl oxygen atom. The radius of the hydroxyl oxygen atom in GLYCAM2000a (1.961 Å) is larger than the OPLS value of 1.7210 Å. The OPLS van der Waals radii are implemented in the AMBER parameter sets as well as GLYCAM06 and here in GLYCAM06-LP.[22,31] The larger van der Waals radius as well as the smaller well depth (0.14 versus 0.21 × 10^{-3} kcal·mol^{-1}) both contributed to the expansion and elongation of hydrogen bonds observed in GLYCAM2000a versus 06 and 06-LP, Tables 3 and 4.

The different partial charge arrangement in each of these force fields would be expected to have a relatively small effect on the overall change in cell dimensions since each charge set was derived from the same set of conformers for the ensemble with fitting to the HF/6-31G(d) MEP and RESP weighting of 0.01. In order to observe the sensitivity purely due to electrostatics, the OPLS hydroxyl oxygen atom van der Waals parameters were implemented in GLYCAM2000a to yield GLYCAM2000b, Table 3. Again, GLYCAM06-LP yields the smallest average unit-cell deviation and reduces the deviation by over 20% from that reported for the GROMOS and HGFB force fields, Table 3.[34] Despite being fit to the same electrostatic potential, the addition of lone-pairs improves the hydrogen bonding interactions over the

Table 3. Effect of Different Force Field Parameters in GLYCAM and Pressure Scaling on the Crystallographic Cell Dimensions for α-D-Glc*p* and the Methyl Glycosides of α-D-Glc*p*, β-D-Gal*p*, and α-D-Man*p*

			cell dimensions (Å)			deviation (%)			
		pressure scaling[a]	*A*	*B*	*C*	Δ*A*	Δ*B*	Δ*C*	mean |Δ|
α-D-Glc*p*									
	2000a	a	21.25	29.86	23.73	4.3	0.5	19.2	8.0
		i	22.14	32.28	21.63				8.7
	2000b	a	20.74	29.70	20.57	1.8	0.0	3.4	1.7
		i	20.68	30.16	20.21				1.5
	06	a	21.04	30.14	19.65	3.3	1.5	−1.3	2.0
		i	20.65	30.10	20.17				1.4
	06-LP	a	20.80	30.03	19.76	2.1	1.1	−0.7	1.3
		i	20.57	29.99	20.10				1.0
	GROMOS[34]	a				−0.8	−3.0	−0.9	1.6
	CHARMMHGFB[34]	a				−3.9	0.8	−0.6	1.8
expt[48]			20.37	29.70	19.90				
α-D-Glc*p*-OMe									
	2000a	a	24.73[b]	32.48	20.00	9.3	9.9	−5.3	8.2
		i	23.76	31.05	22.19				5.0
	06	a	22.91	21.19	20.55	1.3	1.6	−2.7	1.8
		i	22.69	29.65	30.02				0.3
	06-LP	a	22.70	30.40	20.25	0.3	2.8	−4.1	2.4
		i	22.64	29.58	21.14				0.1
expt[47]			22.62	29.56	21.12				
β-D-Gal*p*-OMe									
	2000a	a	33.95	16.52	28.96	9.1	−3.2	10.3	7.5
		i	32.94	18.07	27.80				5.8
	06	a	32.05	16.87	26.09	3.0	−1.1	−0.6	1.6
		i	31.30	17.17	26.42				0.6
	06-LP	a	31.89	16.69	26.26	2.5	−2.2	0.0	1.6
		i	31.21	17.12	26.34				0.3
expt[46]			31.12	17.07	26.26				
α-D-Man*p*-OMe									
	2000a	a	19.81	38.91	21.78	5.1	4.4	8.3	5.9
		i	21.34	39.54	21.34				6.1
	06	a	19.28	37.22	20.17	2.2	−0.1	0.3	0.9
		i	18.99	37.52	20.25				0.7
	06-LP	a	19.16	37.03	20.07	1.6	−0.6	−0.2	0.8
		i	18.89	37.32	20.14				0.2
expt[47]			18.86	37.26	20.11				

[a] Isotropic, i; anisotropic, a. [b] Standard deviations were all within 0.15% of the mean.

more elongated hydrogen bonds found in the models sans lone-pairs. Within each model, there is a weak correlation between the magnitude of the standard deviation and the difference between the calculated and experimental hydrogen bonds, with the smallest deviation of 0.11 Å corresponding to simulated hydrogen bonds that are on average within 0.1 Å of the experimental value. This is a reassuring occurrence in all the models, that they do not adopt the incorrect minimum but are fluctuating over several low energy states. Overall, GLYCAM06-LP has the lowest standard deviation and yields the best reproduction of the hydrogen-bonding environment in the crystal.

Previous studies have examined the ability of carbohydrate force fields to reproduce solvent−solute properties, revealing that GLYCAM2000a underestimates the hydrogen bonding interaction between the hydroxyl groups in the pyranoside and the TIP3P water model.[57] Simulation of the monohydrate glycoside crystal structure required the inclusion of a water molecule, which directly assessed the compatibility of the carbohydrate force field with the chosen solvent model. Here, the TIP series of models was implemented, with the smallest deviation from experiment observed when GLYCAM06-LP was utilized with the TIP5P water model, Table 5. Notably, GLYCAM2000b and GLYCAM06 yield the smallest distortion of the cell dimensions when the TIP5P model is implemented, with TIP3P being the worst, although still in good agreement with the experimental values. Therefore, it is clear not only that under these conditions TIP5P performs better than TIP3P and TIP4P but also that the inclusion of lone-pairs in the carbohydrate force field results in a substantial improvement in the model as well.

In order to assess the performance of the asymmetric LP model, crystal-phase MD simulations were performed on a small peptide, the Ala−Ala (ala_2) zwitterion shown in Figure 8. The *I*4 space group, inherent to the crystal structure of the ala_2 zwitterion, allows MD simulation with both isotropic

Table 4. Effect of Different Force Field Parameters and Pressure Scaling on Selected Intermolecular Hydrogen Bonds (Å) in α-D-Glc*p*, β-D-Gal*p*, and α-D-Man*p*

			interatomic distances		
α-D-Glc*p*	force field	pressure scaling[a]	O2···H3O	O3···H6O	O6···H2O
	GLYCAM 2000a	a	4.05 ± 0.31	2.44 ± 0.36	2.36 ± 0.23
		i	4.08 ± 0.39	2.56 ± 0.65	2.33 ± 0.25
	GLYCAM 06	a	2.13 ± 0.29	1.91 ± 0.18	1.79 ± 0.12
		i	2.05 ± 0.25	1.88 ± 0.16	1.79 ± 0.11
	GLYCAM 06-LP	a	1.97 ± 0.22	1.87 ± 0.14	1.79 ± 0.11
		i	1.91 ± 0.18	1.86 ± 0.14	1.77 ± 0.11
expt[47]			1.770	1.772	1.738
β-D-Gal*p*	force field	pressure scaling[a]	O2····H6O	O6···H4O	O4···H2O
	GLYCAM 2000a	a	4.03 ± 0.87	3.06 ± 0.68	2.39 ± 0.24
		i	4.38 ± 1.33	3.90 ± 1.34	2.98 ± 0.76
	GLYCAM 06	a	3.07 ± 0.47	1.84 ± 0.13	2.03 ± 0.20
		i	2.81 ± 0.72	1.83 ± 0.13	1.97 ± 0.19
	GLYCAM 06-LP	a	2.40 ± 0.44	1.84 ± 0.12	1.98 ± 0.17
		i	2.03 ± 0.27	1.85 ± 0.12	1.89 ± 0.14
expt[46]			1.860	1.739	1.773
α-D-Man*p*	force field	pressure scaling[a]	O3···H6O	O4···H3O	O5···H4O
	GLYCAM 2000a	a	2.32 ± 0.21	2.19 ± 0.17	2.41 ± 0.25
		i	2.35 ± 0.22	2.16 ± 0.15	2.39 ± 0.24
	GLYCAM 06	a	1.92 ± 0.17	1.82 ± 0.11	2.10 ± 0.24
		i	1.90 ± 0.16	1.82 ± 0.12	2.19 ± 0.26
	GLYCAM 06-LP	a	2.00 ± 0.21	1.85 ± 0.12	2.07 ± 0.20
		i	1.96 ± 0.19	1.86 ± 0.13	2.15 ± 0.21
expt[47]			1.917	1.810	2.052

and anisotropic pressure scaling. The deviation of the crystal cell dimensions was similar with all models if isotropic scaling was employed, Table 6. Notably, the asymmetric LP model presented here (LP) yielded the lowest deviation, while the 02EP model was the next best model. Inclusion of polarization into the PARM99 parameter set, PARM02, actually resulted in a worse reproduction of the crystal unit cell dimensions. All models implemented here reproduce the interresidue hydrogen bond distances and angles to within 5% of the experimental values, with PARM02EP yielding the closest agreement, with ≤0.01 Å average deviation in the hydrogen bond distances. Anisotropic pressure scaling on the ala_2 zwitterion had only a subtle, worsening effect on the reproduction of the experimental cell dimensions and hydrogen bond distances, which is in contrast to the large effect anisotropic scaling had on the crystal simulations of monosaccharides.

Inclusion of lone-pairs into a pre-existing molecular mechanics force field may require subsequent refitting of the torsion terms. The partial atomic charges in GLYCAM06-LP were fit to the same MEP as in GLYCAM06, minimizing the impact of the LP-model on existing rotation potentials. This is illustrated clearly for the rotational profiles of the C−O−C−O torsion angle, which is common to all oligosaccharides, and exemplified by axial and equatorial tetrahydro-2-methoxy-2H-pyran, corresponding to α- and β-linkages, respectively, Figure 9. Here, addition of LPs does not have a substantial impact, and the shape of torsional energy curves retains the original optimized shape obtained without LPs.

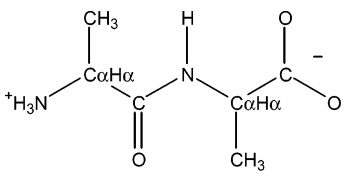

Figure 8. Structure of the Ala−Ala (ala_2) zwitterion.

A highly sensitive measure of the balance between internal rotational energies and external solvent influences is provided by the rotamer population distribution for the exocyclic C5−C6 bond in hexopyranoses. In order to examine the robustness of this LP model, a 10 ns condensed phase MD simulations was performed in conjunction with the TIP5P.[59] Rotamers of the primary alcohol group are populated to varying extents in different monosaccharides, as determined by NMR spectroscopy. The three different rotamers that are populated are defined by the gauche and trans orientation of both the O5−C5−O6−C6 and C4−C5−O6−C6 angles, respectively. All three rotamers are populated for α-D-Gal*p*OMe, with experimentally determined populations *gg*:*gt*:*tg* of 14:47:39,[60] 16:75:9,[61] and 21:61:18,[62] and are reproduced with GLYCAM06 (100 ns) at 8:75:18[30] and here with GLYCAM06-LP at 13:81:6. When 10 ns is too short to ensure statistical convergence for this rotation, longer

Table 5. Effect of Force Field and Water Model on the Crystallographic Unit Cell Parameters and Intermolecular Hydrogen Bond Geometries of α-D-Glc*p*·H_2O

		water model		
force field GLYCAM2000a	expt[49]	TIP3P	TIP4P-EW[58]	TIP5P-EW[59]
A[a]	17.61	18.90 ± 0.03	18.76 ± 0.02	18.73 ± 0.02
B[a]	20.34	21.83 ± 0.04	21.68 ± 0.02	21.64 ± 0.02
C[a]	19.42	20.84 ± 0.04	20.69 ± 0.02	20.66 ± 0.02
Δ*A*, *B*, and *C* (%)[b]		7.33	6.57	6.39
mean |Δ| HB_{dist} (Å)		0.49 ± 0.18	0.36 ± 0.14	4.17 ± 0.73
mean |Δ%| HB_{dist} (%)		23.68	16.97	203.46
mean |Δ| HB_{angle} (°)		72 ± 34	15 ± 17	29 ± 25
mean |Δ%| HB_{angle}		44.25	9.54	16.50

		water model		
force field GLYCAM2000b	expt[49]	TIP3P	TIP4P-EW[58]	TIP5P-EW[59]
A		17.77 ± 0.01	17.74 ± 0.01	17.66 ± 0.01
B		20.54 ± 0.01	20.49 ± 0.01	20.40 ± 0.01
C		19.60 ± 0.01	19.56 ± 0.01	19.47 ± 0.01
Δ*A*, *B*, and *C*		0.96	0.75	0.29
mean |Δ| HB_{dist}		0.17 ± 0.09	0.19 ± 0.06	0.20 ± 0.11
mean |Δ| HB_{dist}		7.78	8.38	9.17
mean |Δ| HB_{angle}		8 ± 12	6 ± 10	7 ± 12
mean |Δ| HB_{angle}		4.70	3.78	4.68

		water model		
force field GLYCAM06	expt[49]	TIP3P	TIP4P-EW[58]	TIP5P-EW[59]
A		17.75 ± 0.01	17.72 ± 0.01	17.62 ± 0.01
B		20.50 ± 0.01	20.48 ± 0.01	20.36 ± 0.01
C		19.57 ± 0.01	19.55 ± 0.01	19.44 ± 0.01
Δ*A*, *B*, and *C*		0.81	0.67	0.11
mean |Δ| HB_{dist}		0.18 ± 0.06	0.17 ± 0.07	0.16 ± 0.08
mean |Δ| HB_{dist}		8.41	8.00	7.55
mean |Δ| HB_{angle}		7 ± 11	6 ± 10	6 ± 10
mean |Δ| HB_{angle}		4.42	3.81	3.56

		water model		
force field GLYCAM06-LP	expt[49]	TIP3P	TIP4P-EW[58]	TIP5P-EW[59]
A		17.62 ± 0.01	17.67 ± 0.01	17.61 ± 0.01
B		20.36 ± 0.01	20.42 ± 0.01	20.35 ± 0.01
C		19.44 ± 0.01	19.49 ± 0.01	19.42 ± 0.01
Δ*A*, *B*, and *C*		0.11	0.39	0.02
mean |Δ| HB_{dist}		0.19 ± 0.07	0.17 ± 0.09	0.16 ± 0.09
mean |Δ| HB_{dist}		8.68	8.06	7.53
mean |Δ| HB_{angle}		7 ± 12	7 ± 11	7 ± 11
mean |Δ| HB_{angle}		4.60	4.13	4.56

[a] In Å. [b] Isotropic pressure scaling.

simulations would be required in order to determine whether this torsion term should be refit. The presence of lone-pairs on both solvent and solute is likely to be particularly influential in modeling dynamic processes, such as conformational lifetimes, bound water occupancies, diffusion rates, and autocorrelation times. It is hoped that the present model will be useful in providing further insight into these phenomena.

Conclusion

Utilizing the quantum mechanical MEP to determine the distance of the lone-pairs from the respective oxygen atoms leads to a description of the molecular electrostatics that is consistent with the currently available TIP5P water model for the hydroxyl and ether type oxygen atoms. The aforementioned sp^3-type oxygen atoms each have a LP−O distance of 0.7 Å and a charge of zero on the oxygen atom.

Table 6. Effect of Force Field Parameter Set on the Crystallographic Unit Cell Dimensions (Å) for the ala_2 Zwitterion[53]

	isotropic scaling[a]						
AMBER parameter set	*A*	*B*	*C*	Δ*A*	Δ*B*	Δ*C*	mean |Δ|
expt	35.97	35.97	20.62				
94	35.47	35.47	20.33	−0.50	−0.50	−0.28	0.43
96	35.50	35.50	20.34	−0.47	−0.47	−0.27	0.41
99	35.52	35.52	20.36	−0.45	−0.45	−0.26	0.39
02	35.44	35.44	20.31	−0.53	−0.53	−0.30	0.45
02EP	35.59	35.59	20.40	−0.38	−0.38	−0.22	0.33
GLYCAM06-LP	36.22	36.22	20.77	0.25	0.25	0.15	0.22
	anisotropic scaling[a]						
AMBER parameter set	*A*	*B*	*C*	Δ*A*	Δ*B*	Δ*C*	mean |Δ|
94	35.63	35.63	20.08	−0.34	−0.34	−0.54	0.41
96	35.64	35.64	20.12	−0.33	−0.33	−0.49	0.38
99	35.64	35.64	20.17	−0.33	−0.33	−0.45	0.38
02	36.17	36.16	19.24	0.20	0.19	−1.38	0.59
02EP	37.79	35.10	19.54	1.82	−0.87	−1.07	1.25
GLYCAM06-LP	35.89	35.89	21.23	−0.08	−0.08	0.61	0.26

[a] All simulations with isotropic pressure scaling resulted in standard deviations of 0.01 Å in cell dimensions, while the maximum increase observed with anisotropic scaling was 0.05 Å.

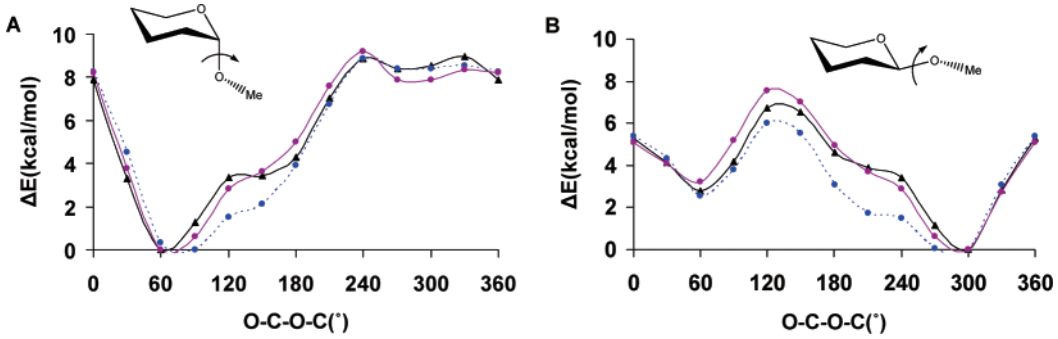

Figure 9. Rotation around the O−C−O−C angle in axial (A) and equatorial (B) tetrahydro-2-methoxy-2*H*-pyran determined at the B3LYP/6-31++G(2d,2p) level (black triangles), GLYCAM06 (purple circles), and single point calculations with GLYCAM06-LP (dashed line, blue circles).

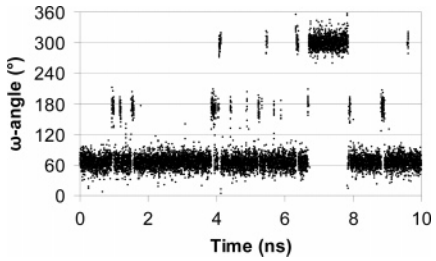

Figure 10. Rotation around the O5−C5−C6−O6 ω-angle in α-D-Gal*p*OMe.

A shorter LP−O distance of 0.3 Å for ketones and an asymmetric LP arrangement, with LP1−O and LP2−O distances of 0.7 and 0.3 Å, for amides yielded the most optimal fit between the quantum mechanical and classical MEP. For each fitting, constraining the oxygen as well as the aliphatic hydrogen atoms to zero charge led to a robust partial charge set with very similar charges for the various atom types in similar environments, i.e., each lone-pair in a secondary alcohol group of α-D-Glc*p* has a partial charge of −0.2 *e*. Inclusion of the new lone-pair model with TIP5P consistently increased the accuracy for MM energy minimized geometries of complexes over those of the non-LP containing model with TIP3P. In addition, the crystal MD simulations clearly illustrate the improved reproduction of electrostatic interactions when LPs are included. With respect to future applications, including LPs may be advantageous in the examination of ligand−receptor complexes, in which water molecules mediate the hydrogen bonding interactions between the ligand and receptor.[63] In addition, it is expected that GLYCAM06-LP will display improved bulk properties such as diffusion rates, rotational correlation times, and radial distribution functions. Extension to a polarizable lone-pair model is currently underway.

Acknowledgment. The authors thank the National Institutes of Health (RR05357 and GM55230) for funding.

References

(1) Leach, A. R. *Molecular Modelling principles and applications*; Addison Wesley Longman Limited: Essex, 1996.

(2) Smit, P. H.; Derissen, J. L.; van Duijneveldt, F. B. *Mol. Phys.* **1979**, *37*, 521.

(3) Scrocco, E.; Tomasi, J. Electronic molecular structure, reactivity and intermolecular forces: An Euristic interpretation by means of electrostatic molecular potentials. In *Advances in Quantum Chemistry;* Löwdin, P., Ed.; 1978; Vol. 11, pp 116−193.

(4) Cox, S. R.; Williams, D. E. *J. Comput. Chem.* **1981**, *2*, 304−323.

(5) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833−1840.

(6) Besler, B. H.; Mertz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431−439.

(7) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361−373.

(8) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894−905.

(9) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129−145.

(10) Woods, R. J.; Khalil, M.; Pell, W.; Moffat, S. H.; Smith, V. H., Jr. *J. Comput. Chem.* **1990**, *11*, 297−310.

(11) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D. M.; Seibel, G. L.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1−41.

(12) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765−784.

(13) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230−252.

(14) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Wang, J.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 7*; University of California: San Francisco, 2002.

(15) Dixon, R. W.; Kollman, P. A. *J. Comput. Chem.* **1997**, *18*, 1632−1646.

(16) 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*, 5179−5197.

(17) Wang, Z. X.; Zhang, W.; Wu, C.; Lei, H. X.; Cieplak, P.; Duan, Y. *J. Comput. Chem.* **2006**, *27*, 994−*994*.

(18) Malcolm, N. O. J.; Popelier, P. L. A. *Faraday Discuss.* **2003**, *124*, 353−363.

(19) Bader, R. F. W.; Matta, C. F. *J. Phys. Chem. A* **2004**, *108*, 8385−8394.

(20) Popelier, P. L. A. *Chem. Phys. Lett.* **1994**, *228*, 160−164.

(21) Harder, E.; Anisimov, V. M.; Vorobyov, I. V.; Lopes, P. E. M.; Noskov, S. Y.; MacKerell, A. D.; Roux, B. *J. Chem. Theory Comput.* **2006**, *2*, 1587−1597.

(22) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638−6646.

(23) Mackerell, A. D.; Wiorkiewiczkuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.

(24) Mahoney, M. W.; Jorgensen, W. L. *J. Chem. Phys.* **2000**, *112*, 8910−8922.

(25) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(26) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. A* **2004**, *108*, 621−627.

(27) Imperiali, B. *Acc. Chem. Res.* **1997**, *30*, 452−459.

(28) Coltart, D. M.; Royyuru, A. K.; Williams, L. J.; Glunz, P. W.; Sames, D.; kuduk, S. D.; Schwarz, J. B.; Chen, X.-T.; Danishefsky, S. J.; Live, D. H. *J. Am. Chem. Soc.* **2002**, *124*, 9833 -9844.

(29) Woods, R. J.; Dwek, R. A.; Edge, C. J. *J. Phys. Chem.* **1995**, *99*, 3832−3846.

(30) Kirschner, K. N.; Yongye, A.; Tschampel, S. M.; Daniels, C.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2007**, Submitted for publication.

(31) Kirschner, K. N.; Woods, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10541−10545.

(32) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.

(33) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X. X.; Murphy, R. B.; Zhou, R. H.; Halgren, T. A. *J. Comput. Chem.* **2002**, *23*, 1515−1531.

(34) Kouwijzer, M. L. C. E.; van Eijck, B. P.; Kroes, S. J.; Kroon, J. *J. Comput. Chem.* **1993**, *14*, 1281−1289.

(35) Woods, R. J.; Chapelle, R. J. *J. Mol. Struct.* **2000**, *527*, 149−156.

(36) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98, (Revision A.9)*; Gaussian, Inc.: Pittsburgh, PA, 1998.

(37) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200−206.

(38) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, 37, 785−789.

(39) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(40) Roothan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69−89.

(41) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358−1371.

(42) Moller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618−622.

(43) Cieplak, P.; Caldwell, J.; Kollman, P. A. *J. Comput. Chem.* **2001**, *22*, 1048−1057.

(44) Kirschner, K. N.; Woods, R. J. *J. Phys. Chem. A* **2001**, *105*, 4150−4155.

(45) Tschampel, S. M.; Woods, R. J. *J. Phys. Chem. A* **2003**, *107*, 9175−9181.

(46) Takagi, S.; Jeffrey, G. A. *Acta Crystallogr., Sect. B: Struct. Sci.* **1979**, *35*, 902−906.

(47) Jeffrey, G. A.; McMullan, R. K.; Takagi, S. *Acta Crystallogr., Sect. B: Struct. Sci.* **1977**, *33*, 728−737.

(48) Brown, G. M.; Levy, H. A. *Acta Crystallogr., Sect. B: Struct. Sci.* **1979**, *35*, 656−659.

(49) Hough, E.; Neidle, S.; Rogers, D.; Troughton, P. G. H. *Acta Crystallogr., Sect. B: Struct. Sci.* **1973**, *B 29*, 365−367.

(50) Mo, F.; Jensen, L. H. *Acta Crystallogr., Sect. B: Struct. Sci.* **1975**, *31*, 2867−2873.

(51) Basma, M.; Sundara, S.; Calgan, D.; Vernali, T.; Woods, R. J. *J. Comput. Chem.* **2001**, *22*, 1125−1137.

(52) Francl, M. M.; Carey, C.; Chirlian, L. E.; Gange, D. M. *J. Comput. Chem.* **1996**, *17*, 367−383.

(53) Fletterick, R.; Tsai, C. C.; Hughes, R. E. *J. Phys. Chem.* **1971**, *75*, 918-&.

(54) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(55) Darden, T. A.; York, D.; Pederson, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(56) Price, D. J.; Brooks, C. L. *J. Chem. Phys.* **2004**, *121*, 10096−10103.

(57) Corzana, F.; Motawia, M. S.; Du Penhoat, C. H.; Perez, S.; Tschampel, S. M.; Woods, R. J.; Engelsen, S. B. *J. Comput. Chem.* **2004**, *25*, 573−586.

(58) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665−9678.

(59) Rick, S. W. *J. Chem. Phys.* **2004**, *120*, 6085−6093.

(60) Bock, K.; Duus, J. O. *J. Carbohydr. Chem.* **1994**, *13*, 513−543.

(61) Nishida, Y.; Ohrui, H.; Meguro, H. *Tetrahedron Lett.* **1984**, *25*, 1575−1578.

(62) Nishida, Y.; Hori, H.; Ohrui, H.; Meguro, H. *J. Carbohydr. Chem.* **1988**, *7*, 239−250.

(63) Loris, R.; Stas, P. P. G.; Wyns, L. *J. Biol. Chem.* **1994**, *269*, 26722−26733.

CT700046J