Published in final edited form as:

Biopolymers. 2013 October ; 99(10): 724–738. doi:10.1002/bip.22286.

Polarizable Empirical Force Field for Acyclic Poly-Alcohols Based on the Classical Drude Oscillator

Xibing He, Pedro E. M. Lopes, and Alexander D. MacKerell Jr

Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, 20 Penn Street, Baltimore, MD 21201

Xibing He: xibing@outerbanks.umaryland.edu; Pedro E. M. Lopes: lopes@outerbanks.umaryland.edu; Alexander D. MacKerell: alex@outerbanks.umaryland.edu

Abstract

A polarizable empirical force field for acyclic polyalcohols based on the classical Drude oscillator is presented. The model is optimized with an emphasis on the transferability of the developed parameters among molecules of different sizes in this series and on the condensed-phase properties validated against experimental data. The importance of the explicit treatment of electronic polarizability in empirical force fields is demonstrated in the cases of this series of molecules with vicinal hydroxyl groups that can form cooperative intra- and intermolecular hydrogen bonds. Compared to the CHARMM additive force field, improved treatment of the electrostatic interactions avoids overestimation of the gas-phase dipole moments, results in significant improvement in the treatment of the conformational energies, and leads to the correct balance of intra- and intermolecular hydrogen bonding of glycerol as evidenced by calculated heat of vaporization being in excellent agreement with experiment. Computed condensed phase data, including crystal lattice parameters and volumes and densities of aqueous solutions are in better agreement with experimental data as compared to the corresponding additive model. Such improvements are anticipated to significantly improve the treatment of polymers in general, including biological macromolecules.

Keywords

CHARMM; ethylene glycol; glycerol; carbohydrate; monosaccharide

INTRODUCTION

Broadly speaking, polyols (polyhydric alcohols, polyalcohols) are a class of organic compounds containing multiple hydroxyl groups. In a narrow sense, especially in food chemistry, the word “polyol” refers to “sugar alcohol”, a hydrogenated form of carbohydrate in which each carbon atom carries an OH group (Figure 1). This special structural feature gives polyols unique properties and, accordingly, a wide range of applications. Particularly relevant is the ability to form intramolecular hydrogen bonds (HBs) between the neighboring hydroxyl moieties, and intermolecular HBs originating three-dimensional HB networks. Because of the low ratio of non-polar groups (methylene and methine groups) to polar hydroxyl groups, ethylene glycol (abbreviated as EG, ethane-1,2-diol) and glycerol (propane-1,2,3-triol), the two smallest polyols, are considered as water analogs and are

Correspondence to: A. D. MacKerell, Jr.: alex@outerbanks.umaryland.edu, 1-410-706-7442.

SUPPORTING INFORMATION

Additional supporting information to this article is available, including python scripts to calculate the CHARMM Molvib input scripts.

commonly used as solvents or co-solvents. When dissolved in water, polyols can form strong intermolecular HBs with water molecules that are able to compete with the water-water HBs and disrupt the formation of ice crystal lattice at temperatures below the freezing point of water. Their pure liquids or aqueous solutions can exist as supercooled liquids.¹⁻³ In industry they are widely used as solvents, antifreeze agents, coolants, lubricants and precursors of chemical reactions. They are also the most commonly used cryoprotectants in the cryopreservation of biological systems, such as oocytes, zygotes, embryos, tissues and organisms, etc. at low temperatures.⁴⁻⁶

Polyols have also been long used as sweeteners, i.e. dietary sugar substitutes.^{7,8} Their intrinsic sweetness is of the same order of magnitude as that of sucrose and they offer significant advantages. Firstly polyol sweeteners are easily manufactured. More importantly, compared to the side effects of sugar consumption, polyol sweeteners can help in the management of obesity and diabetes because they have lower caloric content and a slower yield rate of glucose.^{7,8} In addition, they can help reduce the incidence of dental caries by inhibiting the fermentation of dietary carbohydrate and the formation of insoluble glucan.^{8,10,11} In particular, the five-carbon sugar alcohol xylitol is being used in many nutritional and medical applications.^{11,12} Polyols have also been tested as carbon sources for electricity generation in microbial fuel cells (MFCs), which is one class of biofuel cells.¹³

It is well known that polyols and sugars stabilize macromolecules, especially proteins, in solution without altering their structures, thereby maintaining functionalities.¹⁴⁻³⁰ This is of prime importance in biological processes, in therapeutics and in diagnostics. However, the mechanism has not been elucidated. Several models have been proposed to explain the experimental observations including Wyman linkage functions, preferential interaction, surface tension increment, scaled particle theory, solvent exchange equilibria models, and excluded volume effects etc.²⁴⁻³¹ These models are useful to varying degrees, but they do not offer comprehensive explanations. For instance, while glycerol and sorbitol stabilize the native state of globular proteins, propylene glycol destabilizes globular proteins.³²⁻³⁴ Another example is that polyols destabilize macromolecules in some specific conditions.^{35,36}

Because of the physical, biological and industrial significance, polyols have been subjected to many studies. Numerous experimental investigations have been carried out, but some of the mechanisms of the interactions between polyols and other molecules are still not well understood at the molecular level, such as the stabilizing effects on proteins aforementioned. Molecular dynamics (MD) simulations using classical force field methods have proven to be very useful in the study of structure, dynamics, and thermodynamics of complex systems at the atomic level, and can shed light on this issue. Many MD simulations have been performed on systems containing polyols.³⁷⁻⁷⁹ Most of these simulations focus on the conformations of polyols in gas, crystal, liquid or glass states, or their impact on water structure in solutions. Very recently, simulation studies were undertaken to investigate the interactions between polyols and peptides/proteins.⁴⁶⁻⁵⁴

In order to investigate specific systems and problems, molecular simulations need to be performed using the appropriate force fields. The quality of the force field is the foundation of the reliability of the simulations. Most current available force fields for polyols are pairwise additive models. While they have been successfully applied in MD simulation studies, additive models have their limitations. For example, the need to overestimate molecular dipoles in additive force fields to properly treat condensed, typically aqueous phases, disallows accurate treatment of the molecules in low polarity environments, such as the interior of lipids, organic solvents or, simply, the gas phase.⁸⁰ Empirical force fields continue to evolve with respect to improvements in their parametrization as well as in the

underlying form of their potential energy function. In the context of the energy function, significant effort has recently been put towards inclusion of explicit treatment of electronic polarizability in the model.^{81–86} The inclusion of electronic polarizability is anticipated to more accurately treat the interactions between molecules in environments of varying polarity. In addition, the explicit treatment of electronic polarizability should allow for more accurate treatment of 1) intramolecular conformational energetics and 2) the balance between intra- and intermolecular interactions. Polyalcohols, in which multiple intramolecular HBs are accessible and in which those intramolecular HBs are anticipated to compete with intermolecular HBs, are a class of molecules for which such considerations may be particularly important. As far as we know, there are only two studies about the polarizable models of polyols,^{55, 75} or more specifically, one about EG⁷⁵ and the other about glycerol.⁵⁵

In this study, a polarizable empirical force field is developed for acyclic polyols (sugar alcohols) in which electronic polarizability is introduced via a classical Drude oscillator, as previously described.⁸⁴ Starting from the diol (EG) and triol (glycerol), the model is expanded to tetritols, pentitols and hexitols (Figure 1) and verified by reproducing experimental thermodynamic properties of crystals, bulk liquids, and aqueous solutions. The set of parameters is suitable for applications on longer chain polyols. This study is also the first step in the development of a Drude polarizable force field for carbohydrates, since polyols are often used as prototype molecules for carbohydrates. Together with the developed Drude-based CHARMM polarizable models of water,^{87, 88} alkanes,⁸⁹ alcohols,⁹⁰ amides,⁹¹ ethers,^{92, 93} aromatics,⁹⁴ N containing molecules,⁹⁵ S containing molecules⁹⁶ and nucleic acid bases,⁹⁷ and the Drude polypeptide force field, which is currently under development, we are approaching the goal of developing a comprehensive polarizable force field for biological systems including solvents, co-solvents, co-solutes, and macromolecules like carbohydrates, proteins, lipids and nucleic acids.

METHODS

Empirical force field calculations were performed using the program CHARMM⁹⁸ while quantum mechanical (QM) calculations used the programs Gaussian03⁹⁹ and QChem.¹⁰⁰ Optimized QM geometries for each compound were obtained at the MP2/6-31G* level of the theory. This has been shown to provide molecular geometries consistent with available gas phase experimental data and it has been previously utilized during optimization of the CHARMM all-atom empirical additive force field parameters for various compounds including polyols.^{71,74}

The parameters of the published Drude model for aliphatic alcohols⁹⁰ were used as initial guesses for that of the polyols. The positions of the lone pair sites of the oxygen atom were kept without further adjustment. The atomic charges from the additive CHARMM model, and additive atomic polarizabilities from the work of Miller¹⁰¹ were also used as additional reference data for initial guesses. The partial atomic charges, atomic polarizabilities and Thole-scaling factors were firstly optimized by fitting to the QM electrostatic potential (ESP) maps as previously described.^{90, 102} The ESP grids were placed on concentric nonintersecting Connolly surfaces around the target molecule. Two-part ESP calculations were needed: the “unperturbed” ESP for the molecule alone and the “perturbed” ESPs by placing small point charges (+0.5e) at different location around the molecule to probe the polarization response. The QM ESP maps were calculated using the B3LYP functional^{103–105} with the aug-cc-pVDZ basis set.¹⁰⁶ Optimization of the electrostatic parameters is performed using the FITCHARGE module in the program CHARMM. The obtained electrostatic model was then manually adjusted to reproduce the QM dipole moments of multiple conformations of polyols calculated at the MP2/cc-pVTZ level.¹⁰⁷ Not

only the total dipole moments but also the individual X, Y, and Z components were considered.

QM calculations of intermolecular interaction between EG and water were done by placing single water molecules independently at orientations around EG molecule as shown in Figure 2A. EG was fixed in the MP2/6-31G* optimized geometry and water was constrained to that of the corresponding SWM4-NDP model.^{87,88} The minimum energy interaction distance between each of the hydrogen acceptor-donor pairs was optimized at the MP2/6-31G* level of theory while all other geometrical features were held fixed. Interaction energies were then determined using single point MP2/cc-pVQZ calculations with counterpoise correction for basis set superposition error (BSSE) applied.^{108,109} Interaction energies were the difference between the dimer and the sum of monomer energies. Empirical EG water interactions were performed on a preliminarily relaxed geometry of EG with the water orientation obtained from the corresponding QM calculations. Only the interaction distances were optimized with other geometrical parameters held fixed at their initial values.

QM calculations of the interactions with rare gas atoms were performed using the QM optimized geometries of EG with a rare gas atom (He or Ne) placed at the orientations around EG as illustrated in Figure 2B. A scan in 0.01 Å increments was performed at the MP3/6-311++G(3d,3p) level of theory for each rare gas atom orientation to obtain minimum interaction energy and corresponding distance.¹¹⁰ The resulting interactions were used without BSSE correction.¹¹⁰ Empirical EG – rare gas interactions were performed with preliminarily relaxed geometry of EG and rare gas orientation corresponding to QM calculations.

Equilibrium parameters of bonds and valence angles were optimized targeting the mean values from crystallographic survey of the Cambridge Structural Database¹¹¹ (CSD) of molecules with neighboring polyhydric moieties and supplemented with QM optimized geometries of the model compounds. Internal parameter force constant optimization targeted QM vibrational spectra calculated at the MP2/6-31G* level with a scale factor of 0.9434 applied to vibrational modes to account for limitations in the level of theory.¹¹² The internal coordinate assignment to vibrational frequencies was performed according to the method of Pulay et al.¹¹³ using the MOLVIB utility¹¹⁴ implemented in CHARMM and auxiliary Python scripts (included in the supplementary materials).

Dihedral parameters were optimized by reproducing QM potential energy surfaces (PES). Dihedral angle scans were performed in 10° or 15° increments for the torsion angle with subsequent geometry relaxation performed at the MP2/6-31G* level, which was followed by single-point energy evaluation at the MP2/cc-pVTZ level. The CHARMM PES scans were done following a similar procedure, i.e., the torsion angle of interest was scanned while other geometrical parameters were fully relaxed. The dihedral of interest was restrained to the target values using harmonic restraints of 10,000 kcal/mol/radian with geometry optimization initiated using the QM optimized geometries. Energy minimizations were performed with 500 steps of steepest decent (SD) followed by 500 steps of adapted-basis Newton-Raphson (ABNR) method to a RMS gradient of 10⁻⁵ kcal/(mol*Å). Empirical torsion parameters were optimized to minimize the root-mean-square error (RMSE) between the empirical and QM PES profiles.

The MD simulations of pure solvents, aqueous solutions, and crystals followed the protocols used in the previous CHARMM Drude model parametrization studies⁹⁰⁻⁹⁶ and were described in detail there. The thermodynamic properties including molecular volumes (V_m), heat of vaporizations (ΔH_{vap}), the dielectric constants (ϵ), and the solvation free energies in

water (ΔG_{sol}), the crystal lattice parameters and crystal volumes, and the densities of aqueous solutions were calculated according to corresponding experimental conditions and compared to corresponding experimental data.

RESULTS AND DISCUSSION

Figure 1 shows the structures of acyclic polyol molecules parameterized in this study. They include the diol (EG), the triol (glycerol), tetritols (erythritol, threitol), pentitols (ribitol, arabitol, xylitol) and hexitols (allitol, altritol, glucitol, mannitol, gulitol, iditol, galacticol). Because of the chirality, some polyols (threitol, arabitol, altritol, glucitol, mannitol, gulitol and iditol) have D- and L- mirror images (optical isomers), whereas the others (EG, glycerol, erythritol, ribitol, xylitol, allitol, galacticol) are meso compounds. The parametrization protocol followed those previously used for optimization of the Drude polarizable models for several series of small molecules.^{90–96} Flow diagrams of the parameter optimization process can be found in Ref 96. The first step consists in adjusting the electrostatic parameters, including the partial atomic charges, polarizabilities and Thole scaling factors, using EG and glycerol as the model compounds and the previous Drude parameters for alcohols as initial guesses. Target data for optimization of the electrostatic model include fitting the unperturbed and perturbed QM ESP maps of EG, the QM dipole moments of different conformations of EG and glycerol, the dielectric constant of bulk EG, and the QM intermolecular interactions of EG with water. Then, the bonded parameters for EG and glycerol were adjusted to fit the CSD data, QM optimized geometries, QM vibrational spectra and QM PES. Next, the LJ parameters were optimized to reproduce experimental condensed phase properties of EG and glycerol, supplemented with QM rare gas – EG interactions. Then, the electrostatic and LJ parameters were directly transferred to tetritols, pentitols and hexitols. The missing dihedral parameters were fitted to the QM PES profiles of these polyols based on over 2000 conformations. It should be emphasized that all parameter optimizations were done in an iterative and self-consistent manner so that when one parameter was changed all other parameters were tested and reoptimized if necessary. The optimized parameters were further tested and verified through extended simulations of polyol crystals and aqueous solutions. All results presented here are based on the final model.

Parameter Optimization

EG and Electrostatic Model—EG is the simplest of the polyols and smallest compound having vicinal hydroxyl groups that can form intramolecular HB. It is commonly used as a prototype molecule when designing force fields for sugars. The three successive backbone dihedral angles, H-O-C-C, O-C-C-O, and C-C-O-H, determine the molecular conformation. Assuming typical 3-fold torsion potentials with minima for angles near 60° (gauche clockwise, g+, or simple g), 180° (trans, t), and -60° (gauche counter-clockwise, g- or g'), yields $3^3 = 27$ conformational minima that are stable for the EG molecule. Some of these conformers are equivalent by symmetry, reducing the number of energetically distinct conformers to 10: tGg', gGg', g'Gg', gTg', tTt, tTg, gTg, gGg, tGt, and tGg, where the capital letters (G or T) denote the configuration of the central O-C-C-O dihedral angle.

The electrostatic model was optimized based on EG and then applied to longer polyols. Initial parameters were obtained from the published Drude model for monohydric alcohols.⁹⁰ Partial atomic charges, atomic polarizabilities and Thole-scaling factors were firstly optimized by fitting to the QM ESP maps as previously described.^{90,102} and then manually adjusted because of the following factors. First, the fitted electrostatic parameters vary when different minimum energy conformations of EG are used. Secondly, ranges of atomic polarizabilities and Thole-scaling factors need to be avoided which may lead to “polarization catastrophe” during MD simulations (see Figure S1 of the Supporting

Information). Thirdly, the electrostatic model needs to be adjusted to better reproduce the dipole moments of different conformations of EG and its gas phase – water interactions. In the final model the partial atomic charge of the oxygen atom was moved entirely to the corresponding LP sites.

Figure 3 shows that the calculated dipole moments of the ten minimum energy conformations of EG by the CHARMM Drude model agree well with the QM results, not only for the total dipole moments (Figure 3D), but also for the X (Figure 3A), Y (Figure 3B), and Z (Figure 3C) components. The CHARMM nonpolarizable additive force field systematically overestimates the dipoles, as required to accurately produce the condensed-phase properties.⁸⁰ With the explicit treatment of polarization, such overestimation of the dipoles is no longer necessary. The electrostatic parameters of EG were directly transferred to glycerol and longer polyols except for the charge of methine carbon which was adjusted by adding the charges of a methylene carbon and a methylene hydrogen. The calculated total dipole moments of over 2000 conformations of longer polyols with the Drude model are all in good agreement with QM results, as shown in Figure 4. The RMSE of Drude dipole moments compared to the QM values are: 0.121 Debye for glycerol (468 conformations), 0.263 Debye for tetrityls (72 conformations), 0.290 Debye for pentityls (96 conformations), and 0.273 Debye for hexityls (1727 conformations tested). These results demonstrate the transferability of the electrostatic model.

QM minimum interaction energies and distances for individual water molecules interacting with EG (Figure 2A) were used as additional target data for the optimization of the electrostatic model. The level of agreement between the polarizable model and the QM data is shown in Table I. From Table I it is evident that for the trans conformation (tTt), the level of agreement is very good; for the three lowest energy conformations tGg', gGg', and g'Gg', some the Drude interactions energies are too favorable but were considered acceptable.

The static dielectric constant, which is a measure of the macroscopic zero frequency response to an external electric field, depends on the magnitude of the individual molecular dipoles and their spatial correlation.¹¹⁵ It is an important property that reflects the degree of polarization of individual molecules and their reorientation behavior. The CHARMM additive models have been found to systematically underestimate dielectric constants compared to experimental data due to the inherent limitation of additivity.^{89–96,116} Fitting dielectric constants has been used as an important step in the calibration of CHARMM Drude polarizable models^{89–96} to choose the appropriate scaling of the atomic polarizabilities.^{89–96} The calculated dielectric constant of EG using the Drude model is 42.2 at 298 K, in excellent agreement with the experimental value¹¹⁷ of 40.68.

Internal Parameter Optimization of EG and glycerol—Target data for the equilibrium bond distances and valence angles were obtained from the surveys of the CSD¹¹¹ and from MP2/6–31G* gas-phase optimized geometries. Optimized values for the bonds and valence angles are summarized in Table II. Both the Drude and additive models show very good agreement with the target data, with a slight improvement occurring in the Drude model. Both additive and Drude models target primarily the QM values, and lead to deviations from CSD survey data for some values. For instance, the calculated angle H-OM-CM is ~ 4° smaller than CSD data. The QM values of O-H and C-H bonds are ~0.11 Å longer than that from CSD survey data of 105 records, but agree well with QM values and that of pure EG crystal NOZKES02 in CSD database.

Bond and valence angle force constants were optimized by comparing the empirical vibrational spectra with the QM target data. Results shown in Figures S3 and S4 of the Supporting Information indicate an overall very good agreement in the magnitude and

assignment of each vibrational mode. The dihedral force constants about selected torsions were optimized by fitting the relative energies of different minimum energy conformations and PES of dihedral angles, which will be discussed below.

PES and relative energy profiles of EG conformations—During development of the CHARMM additive all-atom force field for carbohydrates in our group,⁷⁴ difficulties were encountered in simultaneously reproducing the QM relative energies of ten minimum energy conformations and the QM PES profiles for the rotations of H-O-C-C and O-C-C-O torsions for EG. The underlying reason is as follows. Like other pair-wise additive force fields,^{118–120} the dipole moments of molecules in CHARMM are usually overestimated relative to the QM values by 20% to 50%.⁸⁰ This overpolarization relative to the gas phase is required in fixed-charge additive force field models to correctly capture the energetics of condensed-phase polar and aqueous systems.⁸⁰ However, in the case of EG, the resulting negative partial charges on the two oxygens lead to a strong electrostatic repulsion when these two atoms are cis, as occurs in the O-C-C-O dihedral scan and in the two highest-energy conformations in the set of ten minimum-energy conformations. The QM results show a substantially lower barrier, presumably due to electronic polarization, which is absent in the CHARMM additive force field. Figure 5 shows the relative energies of the ten minimum energy conformations and the H-O-C-C and O-C-C-O torsional PES of EG calculated by QM and the CHARMM Drude and additive models. The inherent limitations in the additive model lead to the inability to simultaneously reproduce the relative energies of the minima and two PES. This limitation is largely overcome by the polarizable Drude model, with the largest discrepancy being the relative energies of the two lowest energy minima. The inclusion of electronic polarizability allows for significant improvement in the treatment of the relative energies of a model compound such as EG, a phenomenon that is anticipated to lead to significant improved treatment of conformational energies of macromolecules. Similar results are obtained for glycerol. It is worth mentioning that the discrepancy between the QM and Drude relative energies of the two lowest energy minima of EG may be smaller than shown in Figure 5 if a higher level of QM theory is used. The relative energies of these 10 EG minimum energy conformations have been investigated by several QM studies.^{121–123} These studies indicate that energy difference between tTt and tGg' (or gGg') become smaller when higher level of QM theories were used. For example, the energy difference between tTt and tGg' approximated at CCSD(T)/cc-pVQZ is 0.35 kcal/mol smaller than that calculated from MP2/cc-pVTZ, and the MP2/6-311++G(d,p) calculated difference is 0.38 kcal/mol smaller than MP2/6-31G(d).¹²²

Dihedral parameters of longer polyols were fitted to the QM MP2/cc-pVTZ//MP2/6-31G* PES profiles of glycerol, or QM RIMP2/cc-pVTZ//MP2/6-31G* PES profiles of various tetritol, pentitol and hexitol molecules, with totally over 2000 different conformations sampled. Firstly these parameters were simultaneously fit to conformations for scan points on all of the hexitols, yielding 1727 different conformational energies. Prior to fitting, the parameters for the targeted dihedrals were set to zero, and the total RMSE for all of the conformations was 4.0 kcal/mol. After fitting, which included the 1-, 2-, and 3-fold terms for each dihedral, sampling of force constants from 0 to 3 kcal/mol and sampling phase angles of either 0° or 180°, the RMSE was 0.825 kcal/mol, whereas the RMSE for additive model⁷¹ is 2.5 kcal/mol (Figure 6D). All dihedral parameters needed for the pentitols were covered in hexitol fitting. Using the dihedral parameters fit to the hexitols, the RMSE of the CT-CM-CM-CM dihedral for the pentitols is calculated to be 1.16 kcal/mol, whereas the RMSE for additive model⁷¹ is 1.60 cal/mol (Figure 6C). As for the tetritols, the needed CT-CM-CM-CT dihedral parameter was directly transfer from the CT-CM-CM-CM parameter of hexitols, and the corresponding RMSE is 1.50 kcal/mol (Figure 6B), compared to 3.85 kcal/mol before fitting. The corresponding RMSE for the additive models⁷¹ is 1.87 kcal/mol. The needed CT-CM-CT-OT parameter for glycerol was directly copied from that of CT-CM-

CM-OM (or CM-CM-CT-OT, same values) from the hexitol fitting, and the corresponding RMSE is 0.93 kcal/mol, whereas the RMSE for additive models is 1.65 kcal/mol (Figure 6A). It is worth mentioning that the RMSE between Drude model and MP2/cc-pVTZ for ~460 conformations of glycerol (0.93 kcal/mol) is already comparable to the corresponding RMSE between two QM methods, MP2/6-31G* and MP2/cc-pVTZ, which is 0.77 kcal/mol. This data demonstrates both that the fitting procedure leads to significant improvements in the conformational energetic of the targeted hexitols and also confirms the transferability of the dihedral parameters to the shorter chain linear polyols (Table S1 in the Supporting Information).

Initial Lennard Jones (LJ) parameters for the methylene hydrogen and hydroxyl hydrogen were directly taken from the Drude model of alkanes⁸⁹ and alcohols⁹⁰. The LJ parameters of the carbon and oxygen atoms were then adjusted to reproduce the properties of pure EG and glycerol liquids defined by the experimental molecular volume (V_m) and enthalpy of vaporization (H_{vap}). The methine carbon in the middle position of glycerol shares the same LJ parameters of the methylene carbon at the terminal positions in EG and glycerol. Presented in Table III are the EG and glycerol pure solvent properties for the resulting parameters, along with the previously published additive values.^{71,74} With EG, the Drude values are in good agreement with experiments; satisfactory agreement was also obtained with the additive model. However, the additional of the third vicinal hydroxyl in glycerol leads to poor agreement with the additive model for ΔH_{vap} (an error of -14.2% as compared to experiment). This poor agreement is due to a significant overestimation of the favorable energy of the glycerol monomer in the gas phase, leading to ΔH_{vap} being significantly underestimated ($\Delta H_{vap} = \langle E \rangle_{gas} - \langle E \rangle_{liq} + RT$).⁷¹ The overestimation is due to intramolecular hydrogen bonding between the hydroxyls being too favorable in the gas phase due to overestimation of the hydroxyl partial atomic charges in the additive model. This is evident by calculating ΔH_{vap} using a “gas-phase” monomer energy calculated from conformations of glycerol extracted from the pure solvent simulation, which leads to ΔH_{vap} in good agreement with experiment.⁷¹ The presence of the condensed phase leads to the hydroxyl participating in intermolecular hydrogen bonding, thereby competing for the intramolecular hydrogen bonding resulting in “correct” conformational sampling and the resulting gas phase energy leads to the corresponding improvement in ΔH_{vap} . In contrast, the polarizable Drude results are in excellent agreement with the experimental pure solvent properties when the gas phase energy is obtained from simulations of the monomers. This is due to the ability of the polarizable model to properly treat the gas phase dipole moments, and the corresponding conformational energies, as well as reproduce the non-additive effects occurring in the pure solvent leading to the correct liquid properties. Such capability is not accessible to the additive force field, emphasizing the importance of the explicit treatment of electronic polarization to model systems where significant intramolecular interactions can occur.

The relative values of the individual LJ parameters were also monitored using interactions with rare gases (Figure 2B). The differences between QM and Drude results for the minimum interaction energies and distances as a function of orientation are indicators of the balance of the LJ parameters across different atom types.¹¹⁰ The results from this analysis are included in Table S2 of the Supporting Information. Overall, the agreement is satisfactory.

Table III also shows the solvation free energies (ΔG_{sol}) of EG and glycerol in water calculated using Drude and additive models. The Drude model gives ΔG_{sol} of EG -8.72 kcal/mol, which is 0.58 kcal/mol less favorable than the experimental value (-9.30 kcal/mol), whereas the value calculated from the additive model is 0.70 kcal/mol more favorable than the experimental value. The experimental value of ΔG_{sol} for glycerol (-13.43 kcal/mol) was

determined very recently and was used as one of the target data in the third annual contest of blind prediction of hydration energies in 2009 (SAMPL2).¹²⁴ In this contest, most of the participants gave calculated ΔG_{sol} for glycerol around $-9 \sim -10$ kcal/mol using various theoretical or simulation methods/models.¹²⁵ For the Drude model, the ΔG_{sol} for glycerol is -12.77 kcal/mol, which is only 0.66 kcal/mol less favorable than the experimental value. The ΔG_{sol} for the CHARMM additive model is -10.70 kcal/mol, which is 2.73 kcal/mol less favorable than the experimental value, similar to those values determined as part of the SAMPL2 contest. This further points towards the ability of the Drude model to more accurately reproduce condensed phase properties as compared to additive models.

Parameter Validation

To validate the optimized parameters, MD simulations of crystal and aqueous phase were performed. Both condensed-phase properties and conformational distributions were determined and compared with experimental results.

Crystal MD Simulations—Crystal simulations were performed using model systems obtained from the CSD to validate nonbonded parameters and conformational properties. The compounds included glycerol, tetrityls, pentitols, and hexitols (Table IV). The selected crystals provide a thorough investigation of the crystal lattice parameters and are chosen based on diversity, purity, and resolution. All simulations were performed at room temperature (298 K), in order to match the experimental conditions.

The average sizes of the crystal unit cells from the additive simulations are systematically larger than the experimental values.⁷¹ The average percent error of the simulated unit cell lengths A, B, and C for all of the selected polyols is 2.95%, 1.35%, and 4.55%, respectively, and the average unit cell volume is calculated to have an error of 7.28% (Table IV). This overprediction in the average volume suggests limitations in the ability of a pairwise additive condensed-phase force field designed for liquid simulations to reproduce the crystal phase.⁷¹ The environments surrounding a molecule in the liquid and in the crystal state are considerably different, therefore the inability of the additive force field to quantitatively model the latter environment when parametrized to the former is not entirely surprising, as previously discussed.⁷¹ The Drude model shows significant improvement. The average sizes of the crystal unit cells are not systematically larger than experimental values, with the unit cell parameters of arabitol, ribitol, and mannitol being smaller than experimental data. Secondly, the average percent errors and RMSE of cell lengths and cell volume for all simulated crystals are significantly smaller than that in the additive model (Table IV).

Aqueous-Phase Simulations—To test the behavior of the polyols in aqueous solution, the environment in which the force field is anticipated to be applied primarily, densities for molal solutions of glucitol, mannitol, ribitol, xylitol, galactitol, erythritol, and glycerol were calculated and compared to experimental values. Concentrations varied from dilute (0.1 mol/kg) to highly concentrated (5 mol/kg). For consistency with experimental conditions, glucitol and mannitol were simulated at a temperature of 298 K and pressure of 1 atm, while all other compounds were simulated at 298 K and 3.5 atm. All of the molecular densities calculated by additive model were found to reproduce experimental values within 3% error across the entire 50-fold difference in concentration and at ambient and elevated pressures. Moreover, as the concentration increases there is a trend of decreasing error. The densities calculated by Drude model are consistently in better agreement with experimental values, with errors about 1 % less than additive model (Table V).

Conclusion

This paper presents a polarizable model of acyclic poly-alcohols based on the classical Drude oscillator formalism. The model was optimized to reproduce a variety of gas and condensed-phase target data. Transferability of the optimized parameters was tested via their application in crystal MD simulations and aqueous-phase simulations of triol (glycerol) and various tetrityls, pentitols and hexitols. In the aqueous-phase simulations the experimental densities were reproduced. The quality of the resulting poly-alcohol models in treating various condensed-phase properties supports the transferability of the parameters. Significant care was taken in the concerted optimization of the electrostatic and internal parameters, in particular torsional parameters. Inclusion of polarization results in significant improvement relative to the previously published additive force field, with the Drude model being able to reproduce accurately the QM rotational barriers. Moreover, partial charges were optimized to reproduce values of dipole moments obtained from QM methods. These qualities along with the ability of the polarizable model in reproducing condensed-phase properties in a series of related compounds indicate that the model will be useful for force field-based studies of a variety of acyclic polyalcohols.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank the National Institutes of Health (NIH) for financial support (GM051501, GM070855).

References

1. Sudo S, Shinyashiki N, Yagihara S. *J Mol Liq.* 2001; 90:113–120.
2. Miner, CS.; Dalton, NN. *Glycerol.* Reinhold; New York: 1953.
3. Ryabov YE, Hayashi Y, Gutina A, Feldman Y. *Phys Rev B.* 2003; 67:132202.
4. Fahy GM, Wowk B, Wu J, Paynter S. *Cryobiology.* 2004; 48:22–35. [PubMed: 14969679]
5. Kuwayama M. *Theriogenology.* 2007; 67:73–80. [PubMed: 17055564]
6. Mullen SF, Li M, Li Y, Chen ZJ, Critser JK. *Fertil Steril.* 2008; 89:1812–1825. [PubMed: 17681308]
7. Bar A. *Food Chem.* 1985; 16:231–241.
8. Dills WL Jr. *Annu Rev Nutr.* 1989; 9:161–186. [PubMed: 2669868]
9. den Hartog GJM, Boots AW, Adam-Perrot A, Brouns F, Verkooijen IWC, Weseler AR, Haenen GRMM, Bast A. *Nutrition.* 2010; 26:449–458. [PubMed: 19632091]
10. Bradshaw DJ, Marsh PD. *Caries Res.* 1994; 28:251–256. [PubMed: 8069881]
11. Mäkinen KK, Saag M, Isotupa KP, Olak J, Nömmela R, Söderling E, Mäkinen PL. *Caries Res.* 2005; 39:207–215. [PubMed: 15914983]
12. Mäkinen KK. *Med Hypotheses.* 2000; 54:603–613. [PubMed: 10859647]
13. Catal T, Xu ST, Li KC, Bermek H, Liu H. *Biosens Bioelectron.* 2008; 24:849–854.
14. Back JF, Oakenfull D, Smith MB. *Biochemistry.* 1979; 18:5191–5196. [PubMed: 497177]
15. Gekko KJ. *Biochem.* 1982; 91:1197–1204.
16. Kuznetsova N, Chi SL, Leikin S. *Biochemistry.* 1998; 37:11888–11895. [PubMed: 9718312]
17. Kamiyama T, Sadahide Y, Nogusa Y, Gekko K. *Biochim Biophys Acta.* 1999; 1434:44–57. [PubMed: 10556558]
18. Saunders AJ, Davis-Searles PR, Allen DL, Pielak GJ, Erie DA. *Biopolymers.* 2000; 53:293–307. [PubMed: 10685050]
19. McClements DJ. *Crit Rev Food Sci Nutr.* 2002; 42:417–471. [PubMed: 12401000]

20. Tiwari A, Bhat R. *Biophys Chem.* 2006; 124:90–99. [PubMed: 16828962]
21. Usha R, Maheshwari R, Dhathathreyan A, Ramasami T. *Colloids Surf B.* 2006; 48:101–105.
22. Vagenende V, Yap MGS, Trout BL. *Biochemistry.* 2009; 48:11084–11096. [PubMed: 19817484]
23. Kamal MZ, Ahmad S, Rao NM. *Biophys Chem.* 2011; 156:68–71. [PubMed: 21295397]
24. Gekko K, Timasheff SN. *Biochemistry.* 1981; 20:4667–4676. [PubMed: 7295639]
25. Timasheff SN. *Protein Biochem.* 2002; 41:13473–13482.
26. Liu Y, Bolen DW. *Biochemistry.* 1995; 34:12884–12891. [PubMed: 7548045]
27. Kaushik JK, Bhat R. *J Phys Chem B.* 1998; 102:7058–7066.
28. Xie G, Timasheff SN. *Protein Sci.* 1997; 6:211–221. [PubMed: 9007993]
29. Davis-Searles PR, Saunders AJ, Erie DA, Winzor DJ, Pielak GJ. *Annu Rev Biophys Biomol Struct.* 2001; 30:271–306. [PubMed: 11340061]
30. Gerlsma SY. *Eur J Biochem.* 1970; 14:150–153. [PubMed: 5447430]
31. Knoll D, Hermans J. *J Biol Chem.* 1983; 258:5710–5715. [PubMed: 6853541]
32. Herskovits TT, Gadegbeku B, Jaillet H. *J Biol Chem.* 1970; 245:2588–2598. [PubMed: 5445802]
33. Gekko K, Koga S. *Biochim Biophys Acta.* 1984; 786:151–160.
34. Dib R, Chobert JM, Dalgalarondo M, Haertle T. *Biopolymers.* 1996; 39:23–30. [PubMed: 8924624]
35. Del Vecchio P, Esposito D, Ricchi L, Barone G. *Int J Biol Macromol.* 1999; 24:361–369. [PubMed: 10408643]
36. Singh LR, Poddar NK, Dar TA, Kumar R, Ahmad F. *Life Sci.* 2011; 88:117–125. [PubMed: 21047521]
37. Grigera JR. *J Chem Soc Faraday Trans I.* 1988; 84:2603–2608.
38. Howard E, Grigera JR. *J Chem Soc Faraday Trans.* 1992; 88:437–441.
39. Carlevaro, Manuel; Caffarena, Ernesto R.; Raúl Grigera, J. *Int J Biol Macromol.* 1998; 23:149–155. [PubMed: 9730169]
40. Marañón J, Grigera JR. *J Mol Struct THEOCHEM.* 1998; 431:7–15.
41. Margulies MM, Sixou B, David L, Vigier G, Dolmazon R, Albrand M. *Eur Phys J E.* 2000; 3:55–62.
42. Shah PP, Roberts CJ. *J Phys Chem B.* 2007; 111:4467–4476. [PubMed: 17411089]
43. Lerbret A, Mason PE, Venable RM, Cesàro A, Saboungi M-L, Pastor RW, Brady JW. *Carbohydr Res.* 2009; 344:2229–2235. [PubMed: 19744646]
44. Politi R, Sapir L, Harries D. *J Phys Chem A.* 2009; 113:7548–7555. [PubMed: 19432403]
45. de Waard H, Amani A, Kendrick J, Hinrichs WLJ, Frijlink HW, Anwar J. *J Phys Chem B.* 2010; 114:429–436. [PubMed: 20000804]
46. Liu FF, Ji L, Zhang L, Dong XY, Sun Y. *J Chem Phys.* 2010; 132:225103. [PubMed: 20550422]
47. Gilman-Politi R, Harries D. *J Chem Theory Comput.* 2011; 7:3816–3828.
48. Mehrnejad F, Ghahremanpour MM, Khadem-Maaref M, Doustdar F. *J Chem Phys.* 2011; 134:035104. [PubMed: 21261393]
49. Vagenende V, Trout BL. *Biophys J.* 2012; 103:1354–1362. [PubMed: 22995508]
50. Chen LY. *Biophys Chem.* 2010; 151:178–180. [PubMed: 20573441]
51. Zhang Y, Straub JE. *J Phys Chem B.* 2009; 113:825–830. [PubMed: 19115811]
52. Hénin J, Tajkhorshid E, Schulten K, Chipot C. *Biophys J.* 2008; 94:832–839. [PubMed: 17921212]
53. Curtis JE, Dirama TE, Carri GA, Tobias DJ. *J Phys Chem B.* 2006; 110:22953–22956. [PubMed: 17107124]
54. Dirama TE, Carri GA, Sokolov AP. *J Chem Phys.* 2005; 122:244910. [PubMed: 16035819]
55. Johnson ME, Malardier-Jugrootd C, Head-Gordon T. *Phys Chem Chem Phys.* 2010; 12:393–405. [PubMed: 20023817]
56. Weng LD, Chen C, Zuo JG, Li WZ. *J Phys Chem A.* 2011; 115:4729–4737. [PubMed: 21500852]
57. Egorov AV, Lyubartsev AP, Laaksonen A. *J Phys Chem B.* 2011; 115:14572–14581. [PubMed: 22004353]

58. Busselez R, Lefort R, Ji Q, Affouard FF, Morineau D. *Phys Chem Chem Phys*. 2009; 11:11127–11133. [PubMed: 20024381]
59. Yongye AB, Foley BL, Woods RJ. *J Phys Chem A*. 2008; 112:2634–2639. [PubMed: 18311953]
60. Riggleman RA, de Pablo JJ. *J Chem Phys*. 2008; 128:224504. [PubMed: 18554026]
61. Li DX, Liu BL, Liu YS, Chen CL. *Cryobiology*. 2008; 56:114–119. [PubMed: 18190903]
62. Dashnau JL, Nucci NV, Sharp KA, Vanderkooi JM. *J Phys Chem B*. 2006; 110:13670–13677. [PubMed: 16821896]
63. Dirama TE, Carri GA, Sokolov AP. *J Chem Phys*. 2005; 122:114505. [PubMed: 15836227]
64. Blicke J, Affouard F, Bordat P, Lerbret A, Descamps M. *Chem Phys*. 2005; 317:253–257.
65. Chelli R, Procacci P, Cardini G, Califano S. *Phys Chem Chem Phys*. 1999; 1:879–885.
66. Chelli R, Procacci P, Cardini G, Della Valle RG, Califano S. *Phys Chem Chem Phys*. 1999; 1:871–877.
67. Root LJ, Stillinger FH. *J Chem Phys*. 1989; 90:1200–1208.
68. Szefczyk B, Cordeiro MNDS. *J Phys Chem B*. 2011; 115:3013–3019. [PubMed: 21388132]
69. Lin YS, Hsiao PY, Chieng CC. *J Chem Phys*. 2011; 134:154509. [PubMed: 21513397]
70. Dai JX, Wang L, Sun YX, Wang L, Sun H. *Fluid Phase Equilib*. 2001; 301:137–144.
71. Hatcher ER, Guvench O, MacKerell AD Jr. *J Chem Theory Comput*. 2009; 5:1315–1327. [PubMed: 20160980]
72. Ferrando N, Lachet V, Teuler JM, Boutin A. *J Phys Chem B*. 2009; 113:5985–5995. [PubMed: 19344171]
73. Kyrychenko A, Dyubko TS. *Biophys Chem*. 2008; 136:23–31. [PubMed: 18495323]
74. Guvench O, Greene SN, Kamath G, Brady JW, Venable RM, Pastor RW, Mackerell AD Jr. *J Comput Chem*. 2008; 29:2543–2564. [PubMed: 18470966]
75. Geerke DP, Van Gunsteren WF. *Mol Phys*. 2007; 105:1861–1881.
76. de Oliveira OV, Freitas LCG. *J Mol Struct THEOCHEM*. 2005; 728:179–187.
77. Stubbs JM, Potoff JJ, Siepmann JI. *J Phys Chem B*. 2004; 108:17596–17605.
78. Saiz L, Padró JA, Guàrdia E. *J Chem Phys*. 2001; 114:3187–3199.
79. Widmalmt G, Pastor RW. *J Chem Soc Faraday Trans*. 1992; 88:1747–1754.
80. Mackerell AD Jr. *J Comput Chem*. 2004; 25:1584–1604. [PubMed: 15264253]
81. Halgren TA, Damm W. *Curr Opin Struct Biol*. 2001; 11:236–242. [PubMed: 11297934]
82. Warsl A, Kato M, Pislakov AV. *J Chem Theory Comput*. 2007; 3:2034–2045.
83. Illingworth CJ, Domene C. *Proc R Soc A*. 2009; 465:1701–1716.
84. Lopes PEM, Roux B, MacKerell AD Jr. *Theor Chem Acc*. 2009; 124:11–28. [PubMed: 20577578]
85. Lucas TR, Bauer BA, Patel S. *Biochim Biophys Acta Biomembr*. 2012; 1818:318–329.
86. Ponomarev SY, Kaminski GA. *J Chem Theory Comput*. 2011; 7:1415–1427. [PubMed: 21743799]
87. Lamoureux G, MacKerell AD Jr, Roux B. *J Chem Phys*. 2003; 119:5185–5197.
88. Lamoureux G, Harder E, Vorobyov IV, Roux B, MacKerell AD Jr. *Chem Phys Lett*. 2006; 418:245–249.
89. Vorobyov IV, Anisimov VM, MacKerell AD Jr. *J Phys Chem B*. 2005; 109:18988–18999. [PubMed: 16853445]
90. Anisimov VM, Vorobyov IV, Roux B, MacKerell AD Jr. *J Chem Theory Comput*. 2007; 3:1927–1946. [PubMed: 18802495]
91. Harder E, Anisimov VM, Whitfield T, MacKerell AD Jr, Roux B. *J Phys Chem B*. 2008; 112:3509–3521. [PubMed: 18302362]
92. Vorobyov I, Anisimov VM, Greene S, Venable RM, Moser A, Pastor RW, MacKerell AD Jr. *J Chem Theory Comput*. 2007; 3:1120–1133.
93. Baker CM, MacKerell AD Jr. *J Mol Model*. 2010; 16:567–576. [PubMed: 19705172]
94. Lopes PEM, Lamoureux G, Roux B, MacKerell AD Jr. *J Phys Chem B*. 2007; 111:2873–2885. [PubMed: 17388420]
95. Lopes PEM, Lamoureux G, MacKerell AD Jr. *J Comput Chem*. 2009; 30:1821–1838. [PubMed: 19090564]

96. Zhu X, Mackerell AD Jr. *J Comput Chem.* 2010; 31:2330–2341. [PubMed: 20575015]
97. Baker CM, Anisimov VM, MacKerell AD Jr. *J Phys Chem B.* 2011; 115:580–596. [PubMed: 21166469]
98. Brooks BR, Brooks CL III, MacKerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M. *J Comput Chem.* 2009; 30:1545–1614. [PubMed: 19444816]
99. Frisch, MJT.; Schlegel, GW.; Scuseria, GE.; Robb, MA.; Cheeseman, JR.; Montgomery, JA., Jr; Vreven, T.; Kudin, KN.; Burant, JC.; Millam, JM.; Iyengar, SS.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, GA.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, JE.; Hratchian, HP.; Cross, JB.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, RE.; Yazyev, O.; Austin, AJ.; Cammi, R.; Pomelli, C.; Ochterski, JW.; Ayala, PY.; Morokuma, K.; Voth, GA.; Salvador, P.; Dannenberg, JJ.; Zakrzewski, VG.; Dapprich, S.; Daniels, AD.; Strain, MC.; Farkas, O.; Malick, DK.; Rabuck, AD.; Raghavachari, K.; Foresman, JB.; Ortiz, JV.; Cui, Q.; Baboul, AG.; Clifford, S.; Cioslowski, J.; Stefanov, BB.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, RL.; Fox, DJ.; Keith, T.; Al-Laham, MA.; Peng, CY.; Nanayakkara, A.; Challacombe, M.; Gill, PMW.; Johnson, B.; Chen, W.; Wong, MW.; Gonzalez, C.; Pople, JA. Gaussian 03, Revision C.02. Gaussian Inc; Wallingford CT: 2004.
100. Shao Y, Molnar LF, Jung Y, Kussmann J, Ochsenfeld C, Brown ST, Gilbert AT, Slipchenko LV, Levchenko SV, O'Neill DP, Distasio RA, Lochan RC, Wang T, Beran GJ, Besley NA, Herbert JM, Lin CY, Van Voorhis T, Chien SH, Sodt A, Steele RP, Rassolov VA, Maslen PE, Korambath PP, Adamson RD, Austin B, Baker J, Byrd EF, Dachsel H, Doerksen RJ, Dreuw A, Dunietz BD, Dutoi AD, Furlani TR, Gwaltney SR, Heyden A, Hirata S, Hsu CP, Kedziora G, Khalliulin RZ, Klunzinger P, Lee AM, Lee MS, Liang W, Lotan I, Nair N, Peters B, Proynov EI, Pieniazek PA, Rhee YM, Ritchie J, Rosta E, Sherrill CD, Simmonett AC, Subotnik JE, Woodcock HL, Zhang W, Bell AT, Chakraborty AK, Chipman DM, Keil FJ, Warshel A, Hehre WJ, Schaefer HF, Kong J, Krylov AI, Gill PM, Head-Gordon M. *Phys Chem Chem Phys.* 2006; 8:3172–3191. [PubMed: 16902710]
101. Miller KJ. *J Am Chem Soc.* 1990; 112:8533–8542.
102. Anisimov VM, Lamoureux G, Vorobyov IV, Huang N, Roux B, MacKerell AD Jr. *J Chem Theory Comput.* 2005; 1:153–168.
103. Lee C, Yang W, Parr RG. *Phys Rev B.* 1988; 37:785–789.
104. Becke AD. *Phys Rev A.* 1988; 38:3098–3100. [PubMed: 9900728]
105. Becke AD. *J Chem Phys.* 1993; 98:5648–5652.
106. Kendall RA, Dunning TH Jr, Harrison RJ. *J Chem Phys.* 1992; 96:6796–6806.
107. Dunning THJ. *J Chem Phys.* 1989; 90:1007–1023.
108. Boys SF, Bernardi F. *Mol Phys.* 1970; 19:553–566.
109. Ransil BJ. *J Chem Phys.* 1961; 34:2109–2118.
110. Yin D, MacKerell AD Jr. *J Comp Chem.* 1998; 19:334–348.
111. Allen FH. *Acta Crystallogr.* 2002; B58:370–379.
112. Scott AP, Radom L. *J Phys Chem.* 1996; 100:16502–16513.
113. Pulay P, Fogarasi G, Pang F, Boggs JE. *J Am Chem Soc.* 1979; 101:2550–2560.
114. Kuczera K, Wiorkiewicz-Kuczera J. MOLVIB program. 1991
115. Kirkwood JG. *J Chem Phys.* 1939; 7:911–919.
116. Whitfield TW, Martyna GJ, Allison S, Bates SP, Vass H, Crain J. *J Phys Chem B.* 2006; 110:3624–3637. [PubMed: 16494418]
117. Sengwa RJ, Sankhla S, Shinyashiki N. *Phys Chem Liquids.* 2010; 48:29–40.
118. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. *J Am Chem Soc.* 1995; 117:5179–5197.
119. Jorgensen WL, Tirado-Rives J. *J Am Chem Soc.* 1988; 110:1657–1666.

120. Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF. J Comput Chem. 2004; 25:1656–1676. [PubMed: 15264259]
121. Cramer CJ, Truhlar DG. J Am Chem Soc. 1994; 116:3892–3900.
122. Guvench O, MacKerell AD Jr. J Phys Chem A. 2006; 110:9934–9939. [PubMed: 16898697]
123. Cheng Y-L, Chen H-Y, Takahashi K. J Phys Chem A. 2011; 115:5641–5653. [PubMed: 21568300]
124. Geballe MT, Skillman AG, Nicholls A, Guthrie JP, Taylor PJ. J Comput Aided Mol Des. 2010; 24:259–279. [PubMed: 20455007]
125. Geballe, MT. 2010. http://www.eyesopen.com/2010_cup_presentations/SAMPL-the-beat-goes-on.pdf

FIGURE 1.

Acyclic polyalcohols. Those compounds not designated with a D are meso compounds. C1 is at the top most position of the carbon chain.

FIGURE 2. Interaction orientations of ethylene glycol with water (A) and rare gases (B). rare-gas atoms are shown in green.

FIGURE 3.

The X (A), Y (B), Z (C), and total (D) dipole moments of 10 minimum energy conformations of ethylene glycol calculated by QM (MP2/cc-pVTZ, red cross symbols), CHARMM Drude polarizable force field (blue solid lines), and CHARMM nonpolarizable additive force field (green dash lines).

FIGURE 4. The calculated dipole moments of different conformations of glycerol (A), tetritols (B), pentitols (C) and hexitols (D).

FIGURE 5.

Relative energies of ethylene glycol calculated by QM, CHARMM Drude polarizable force field, and CHARMM additive force field.

FIGURE 6. QM and MM potential energy scans for glycerol (A), tetritols (B), pentitols (C) and hexitols (D).

Table I

QM and Drude Ethylene Glycol-Water Minimum Interaction Energies and Distances.

Conf.	Direction	E_{QM} (kcal/mol)	E_{Drude} (kcal/mol)	$E_{Drude}-E_{QM}$ (kcal/mol)	R_{QM} (Å)	R_{Drude} (Å)	$R_{Drude}-R_{QM}$ (Å)
tGg'	PR	-5.45	-4.97	0.48	1.96	1.92	-0.04
	LP	-4.46	-4.95	-0.49	2.06	1.82	-0.24
	BIS	-3.46	-3.02	0.44	2.06	1.88	-0.18
g'Gg'	PR	-5.65	-4.89	0.76	1.97	1.93	-0.04
	LP	-4.40	-4.31	0.09	2.05	1.85	-0.20
	BIS	-3.68	-2.63	1.05	2.09	1.91	-0.18
g'Gg'	PR	-4.26	-4.02	0.24	2.03	1.96	-0.07
	LP	-4.74	-4.69	0.05	2.03	1.83	-0.20
	BIS	-4.70	-3.76	0.94	2.03	1.86	-0.17
tTt	PR	-4.90	-4.63	0.27	1.98	1.92	-0.06
	LP	-4.92	-5.04	-0.12	2.04	1.84	-0.20
	BIS	-4.77	-4.58	0.19	2.04	1.85	-0.19

Table II

Internal Geometries of Ethylene Glycol and Glycerol.

Ethylene Glycol									
Bonds	CSD ^a	QM	Additive ^c	Additive-CSD	Additive-QM	Drude ^c	Drude-CSD	Drude-QM	Drude-QM
OT-H	0.97	0.97	0.96	0.00	-0.01	0.96	-0.01	-0.01	-0.01
OT-CT	1.43	1.43	1.43	0.00	0.00	1.43	0.00	0.00	0.00
CT-CT	1.51	1.52	1.54	0.03	0.03	1.52	0.01	0.01	0.00
CT-H	1.09	1.10	1.11	0.02	0.02	1.09	-0.01	-0.01	-0.01
Average				0.01	0.01		0.00	0.00	0.00
RMSE				0.02	0.02		0.01	0.01	0.01
Glycerol									
Bonds	CSD ^b	QM	Additive ^d	Additive-CSD	Additive-QM	Drude ^d	Drude-CSD	Drude-QM	Drude-QM
OT-H	0.85	0.97	0.96	0.11	-0.01	0.96	0.11	-0.01	-0.01
OM-H	0.85	0.98	0.96	0.11	-0.01	0.96	0.11	-0.01	-0.01
OT-CT	1.43	1.43	1.43	0.01	0.01	1.43	0.01	0.00	0.00
OM-CM	1.43	1.43	1.43	0.00	0.00	1.43	0.00	0.00	0.00
CT-CM	1.52	1.52	1.52	0.00	0.00	1.53	0.01	0.01	0.01
CT-H	0.99	1.10	1.11	0.12	0.02	1.08	0.09	-0.01	-0.01
CM-H	0.99	1.10	1.12	0.13	0.02	1.09	0.10	-0.01	-0.01
Average				0.07	0.00		0.06	-0.01	-0.01
RMSE				0.09	0.01		0.08	0.01	0.01

Ethylene Glycol								
Bonds	CSD ^a	QM	Additive ^c	Additive-CSD	Additive-QM	Drude ^c	Drude-CSD	Drude-QM
Angles	CSD ^b	QM	Additive ^d	Additive-CSD	Additive-QM	Drude ^d	Drude-CSD	Drude-QM
H-OT-CT	108.9	106.8	106.9	-2.0	0.1	107.3	-1.6	0.5
H-OM-CM	110.1	106.0	104.9	-5.2	-1.1	106.6	-3.5	0.6
OT-CT-CM	111.5	109.8	110.6	-0.9	0.8	111.2	-0.3	1.4
OM-CM-CT	109.3	108.6	108.8	-0.5	0.3	109.2	-0.1	0.7
OT-CT-H	108.8	109.9	109.0	0.2	-0.8	109.7	0.9	-0.2
OM-CM-H	108.5	109.2	108.3	-0.2	-0.9	108.6	0.1	-0.6
CT-CM-H	108.6	108.8	109.2	0.6	0.4	108.4	-0.2	-0.4
CM-CT-H	109.8	109.4	110.1	0.3	0.6	109.3	-0.5	-0.1
CT-CM-CT	112.4	112.7	112.3	-0.1	-0.5	112.8	0.4	0.1
Average				-0.9	-0.1		-0.5	0.2
RMSE				1.9	0.7		1.3	0.6

^a from crystal NOZKES02.

^b searching using HOCH2CH(OH)CH(OH)R, where R means C or H; 105 records with R factor < 0.5 found from CSD.

^c averaged over 10 minimum energy conformations.

^d averaged over 126 minimum energy conformations. The unit of bond lengths is Å, and the unit of valence angles is degree.

^e CT denotes terminal methylene carbon, CM denotes middle methine carbon, OT is connected to CT, and OM is connected to CM. See Figure S4 of Supporting Information.

Table III

Thermodynamic Properties of Ethylene Glycol and Glycerol.

		V_m		ΔH_{vap}		ΔG_{sol}	
		$\text{\AA}^3$	% Error	kcal/mol	% Error	kcal/mol	Error
Ethylene Glycol	Expt.	92.56 ^a		15.70 ^a		-9.30 ^a	
	Drude	92.13	-0.5%	15.72	+0.1%	-8.72	+0.58
	Additive	95.64 ^a	+3.3% ^a	15.79 ^a	+0.6% ^a	-10.00 ^a	-0.70 ^a
Glycerol	Expt.	121.40 ^b		21.90 ^b		-13.43	
	Drude	119.98	-1.2%	21.67	-1.0%	-12.77	+0.66
	Additive	127.17 ^b	+4.8%	18.80 ^{bc}	-14.2% ^{bc}	-10.70 ^d	+2.73
				22.30 ^{bc}	2.0 % ^{bc}		

^aValues are taken from Ref 74.

^bValues are taken from Ref 71.

^cUpper additive ΔH_{vap} calculated using standard approach with the lower value calculated using gas phase energies required for the ΔH_{vap} calculation determined from conformations of glycerol obtained from the condensed phase simulation.⁷¹

^dValue of ΔG_{sol} was calculated using the additive model of glycerol in Ref 71.

Table IV

Crystal Lattice Parameter and Volumes Calculated from Crystal Simulations.

Compound	CSD ID	Note	A	% error	B	% error	C	% error	volume	% error
allitol					N=6					
	ALITOL01	R factor 0.04	4.71		13.41		6.62		411.16	
		Additive	4.96	5.35	13.81	3.03	6.82	3.14	445.30	8.30
altritol		Drude	4.74	0.64	13.51	0.75	6.67	0.76	420.97	2.39
	JOJZOX	R factor 0.03	4.90		5.18		16.26		409.11	
		Additive	5.61	14.51	5.16	-0.39	16.11	-0.89	446.39	9.11
galactitol		Drude	4.92	0.41	5.20	0.58	16.34	0.49	415.11	1.47
	GALACT	R factor 0.05	8.45		11.50		9.04		808.64	
		Additive	8.66	2.53	11.79	2.45	9.74	7.69	857.58	6.05
glucitol		Drude	8.50	0.59	11.58	0.70	9.10	0.66	825.20	2.05
	GLUCIT01	R factor 0.07	8.68		9.31		9.73		785.86	
		Additive	8.71	0.44	9.48	1.84	10.11	3.89	834.77	6.22
mannitol		Drude	8.76	0.92	9.40	0.97	9.82	0.92	809.9	3.06
	DMANTL07	R factor 0.03	8.69		16.90		5.55		815.40	
		Additive	8.96	3.07	17.22	1.91	5.61	1.03	865.23	6.11
arabitol		Drude	8.67	-0.23	16.85	-0.30	5.53	-0.36	808.54	-0.84
	ARABOL	R factor 0.04	9.21		4.86		15.49		692.85	
		Additive	9.36	1.59	4.95	2.00	15.39	-0.64	713.10	2.92
ribitol		Drude	9.14	-0.76	4.82	-0.82	15.37	-0.77	677.23	-2.25
	RIBITOL	R factor 0.06	8.99		4.95		15.73		694.11	
		Additive	9.31	3.60	5.03	1.64	15.57	-1.04	720.91	3.86
xyitol		Drude	8.98	-0.11	4.94	-0.20	15.71	-0.13	692.12	-0.29
	XYLTOL01	R factor 0.05	8.27		8.90		8.91		655.43	
		Additive	8.37	1.27	9.39	5.47	9.18	3.00	720.74	9.96
		Drude	8.35	0.97	8.99	1.01	9.00	1.01	676.08	3.15
					N=4					

Compound	CSD ID	Note	A	% error	B	% error	C	% error	volume	% error
threitol	PAGDEG	R factor 0.05	10.10		10.10		4.84		427.60	
		Additive	9.72	-3.66	9.72	-3.66	5.85	20.87	478.68	11.95
		Drude	10.19	0.89	10.19	0.89	4.88	0.83	439.18	2.71
					N=3					
glycerol	GLCROL	R factor 0.12	7.00		9.96		6.29		438.54	
		Additive	7.06	0.79	9.88	-0.78	6.82	8.43	474.91	8.30
		Drude	7.05	0.71	10.03	0.70	6.33	0.64	447.96	2.15
Average		Additive		2.95		1.35		4.55		7.28
		Drude		0.40		0.43		0.41		1.36
RMSE		Additive		5.35		2.70		7.76		7.74
		Drude		0.68		0.74		0.70		2.22

Experimental and Calculated Densities at Different Concentrations of Polyols in a Box of Water with 1100 Water Molecules at $T = 298.15$ K and $P = 1$ or 3.5 atm.

Table V

Compound	Molality (mol/kg)	N_{solute}	Expt (g/cc)	Additive (g/cc)	%error	Drude (g/cc)	%error
mannitol ($P=1$ atm)	0.1999	4	1.0092	1.0228	1.35	1.0135	0.43
	0.5998	12	1.0318	1.0414	0.93	1.0369	0.49
	0.8006	16	1.0426	1.0503	0.74	1.0479	0.51
	0.9995	20	1.0517	1.0584	0.64	1.0575	0.55
glucitol ($P=1$ atm)	0.5085	10	1.0117	1.0368	2.48	1.0310	1.91
	1.9508	39	1.0987	1.0921	-0.60	1.0997	0.09
	4.0003	79	1.1606	1.1445	-1.39	1.1665	0.51
	5.9945	119	1.2052	1.1793	-2.15	1.2140	0.73
galacticol ($P=3.5$ atm)	0.0698	2	1.0015	1.0179	1.64	1.0076	0.61
	0.1492	3	1.0065	1.0202	1.36	1.0104	0.39
xyliol ($P=3.5$ atm)	0.1000	2	1.0022	1.0165	1.43	1.0058	0.36
	1.0000	20	1.0427	1.0472	0.43	1.0464	0.35
	2.7000	53	1.1019	1.0925	-0.85	1.1029	0.09
ribitol ($P=3.5$ atm)	0.1042	2	1.0022	1.0164	1.42	1.0059	0.37
	0.5092	10	1.0208	1.0315	1.05	1.0249	0.41
	3.1769	63	1.1121	1.1047	-0.67	1.1158	0.34
erythritol ($P=3.5$ atm)	0.4998	10	1.0142	1.0262	1.18	1.0185	0.42
	1.0000	20	1.0298	1.0382	0.82	1.0341	0.42
	3.0000	60	1.0806	1.0804	-0.02	1.0853	0.43
glycerol ($P=3.5$ atm)	0.5002	10	1.0075	1.0200	1.24	1.0112	0.37
	1.0008	20	1.0171	1.0280	1.07	1.0211	0.40
	3.0035	60	1.0497	1.0563	0.63	1.0533	0.34
	4.9993	100	1.0750	1.0781	0.29	1.0781	0.28
RMSE					1.20		0.58

The experimental values and Additive values are from Refs 71. Drude values are from this work.