Prediction of the *n*-Hexane/Water and 1-Octanol/Water Partition Coefficients for Environmentally Relevant Compounds using Molecular Simulation

Nuno M. Garrido
LSRE Laboratory of Separation and Reaction Engineering, Departamento de Engenharia Química, Faculdade de Engenharia, Universidade do Porto, Rua do Dr. Roberto Frias, 4200 - 465 Porto, Portugal

The Petroleum Institute, Dept. of Chemical Engineering, PO Box 2533, Abu Dhabi, United Arab Emirates

Ioannis G. Economou
The Petroleum Institute, Dept. of Chemical Engineering, PO Box 2533, Abu Dhabi, United Arab Emirates

António J. Queimada, Miguel Jorge, and Eugénia A. Macedo
LSRE Laboratory of Separation and Reaction Engineering, Departamento de Engenharia Química, Faculdade de Engenharia, Universidade do Porto, Rua do Dr. Roberto Frias, 4200 - 465 Porto, Portugal

DOI 10.1002/aic.12718
Published online August 9, 2011 in Wiley Online Library (wileyonlinelibrary.com).

In recent years molecular simulation has emerged as a useful tool to predict physical properties of complex chemical systems. A methodology to estimate the n*-hexane/water and 1-octanol/water partition coefficients of environmentally relevant solutes, namely substituted alkyl-aromatic molecules, chlorobenzenes, polychlorinated biphenyls (PCBs) and polychlorinated diphenyl ethers (PCDEs) using molecular simulation is elucidated here. The partition coefficients are calculated based on the absolute solvation Gibbs energies in each phase which are estimated from molecular dynamics simulations employing the thermodynamic integration approach. Very encouraging results, with average absolute deviations of 0.4 log* P *units are presented. Consequently, this molecular-based approach with a strong physical background can provide reliable prediction of the partition coefficients in different solvent pairs without the* a priori *knowledge of experimental data.* © 2011 American Institute of Chemical Engineers *AIChE J,* 58: 1929–1938, 2012

Introduction

Many important biochemical processes involve the interaction of solutes with both hydrophobic and hydrophilic environments. In this context, it becomes extremely important to understand how compounds partition and interact between two or more different fluid phases. Partition coefficients reflect the solute hydrophobicity, partitioning between different solvents or pharmacokinetic characteristics. They are widely used as a measure of environmental fate and have found several applications in the pharmaceutical (e.g., drug design) or chemical industries (e.g., separation technology).[1]

Undoubtedly, the most widely used partition coefficient is the 1-octanol/water partition coefficient (hence, referred as log P_{OW}). The amphiphilic nature of the 1-octanol molecules, with a polar head group attached to a flexible nonpolar tail, affords them similar characteristics to the main constituents of biological membranes.[2,3] 1-Octanol molecules can also mimic the complex behavior of soil. Thus, log P_{OW} plays an important role in the prediction of solute partitioning in environmental fate and toxicological processes. Furthermore, hydrophobicity is conventionally expressed by the log P_{OW} value, where a positive value of this ratio (for the case of a lipophilic substance) reflects a preference for the organic phase while a negative value (for the case of lipophobic substance) indicates an affinity for water. Hydrophobicity (or lipophilicity, in the opposite sense) is a key descriptor used to assess and model the distribution and transport potential of pollutants in biological and environmental systems. Prediction of drug partitioning and pharmacokinetic characteristics in biological systems can also be quantified by expressions based on this partition coefficient.[4,5] Finally, log P_{OW} may be useful to estimate the solubility of a solute in a solvent[6,7] or used as a measure of the bioconcentration factor.[8,9]

Apart from 1-octanol, other lipophilic phases are also important to assess the potential hazard of pollutant components in the environment. An example is *n*-hexane as a model of an inert and hydrophobic organic phase. *n*-hexane/water partition coefficients are used in environmental and chemical engineering or as inputs of linear solvation energy relationships.[10] These partitioning data are important in the chemical engineering field for a rapid selection of a solvent for liquid-liquid extraction from multicomponent aqueous solution. They are also important when taken together with log P_{OW}, as one can compare the solute-solvent intermolecular

Correspondence concerning this article should be addressed to I. G. Economou at ieconomou@pi.ac.ae.

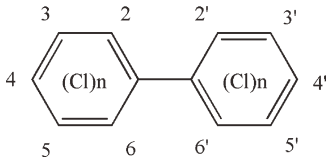

Figure 1. Polychlorinated biphenyl (PCB) structure and atom numbering.

interactions (1-octanol is a polar hydrogen bonding molecule, whereas *n*-hexane is a hydrophobic compound), and extract different structural information by analyzing the behavior of the same chemical in different environments.

Partition coefficients can be measured experimentally by applying, for example, the shake-flask method[11,12] where saturated liquid phases are generated and then a solute sample is quantitatively analyzed (e.g., by high-performance liquid chromatography). Still, this can be a very expensive and time-consuming procedure. In some cases, such as for hydrophobic low aqueous solubility compounds, log *P* values are not easy to determine experimentally, given that direct methods are difficult to apply due to analytical limitations. Moreover, experimental measurements may not even be feasible in cases where the compound of interest is unavailable or has not yet been synthesized (e.g., drug design). For all of these reasons, it is very important to develop accurate predictive methods for partition coefficients. It is worthwhile to notice that for environmental issues, theoretical approaches become an even more important tool because they avoid direct handling of highly toxic, carcinogenic or generally hard-to-handle compounds for which log *P* is required.

The concept of modeling theoretically the partition coefficient of a solute between 1-octanol and water was first introduced in 1964 by Hansch and Fujita.[13] In the beginning, mostly semiempirical approaches based on the sum of fragment contributions or atom-derived group equivalents were proposed.[14] Nowadays, fragment additive schemes remain a standard method to estimate solvation free energies and partition coefficients,[15] but the most common methods to estimate solvation properties are based on quantitative structure–property relationships (QSPR). These are statistically-based techniques that correlate partition coefficients or solvation properties with other calculated or available molecular properties. QSPR methods are considerably fast and applicable to large databases of molecular structures but they also require the input of large multiparameter tables. More importantly, they often lose predictive ability for compounds that are significantly different from their training set, and have, thus, been the object of criticism.[16] Naturally, there is a lack of existing parameters to calculate log P_{OW} for new chemical groups, which is a major drawback of such methods. Linear solvation energy relationships (LSERs) are also used to reproduce log P_{OW} data, as shown in two recent critical reviews.[17,18]

The majority of the previously mentioned empirical methods only describe the overall process of solute distribution and sometimes lack information about the underlying thermodynamic processes. One exception was the prediction with a quite simple and fast model called mobile order and disorder thermodynamics theory presented by Ruelle[19] and applied to a very large database. In another work, Duffy and Jorgensen[20] used simulations based on linear response theory and molecular descriptors to derive empirical relationships for estimating generic log *P* values. Finally, approaches based on free energy perturbation employing continuum models based on electronic structure calculations have also been investigated.[21–23]

In this work, we will explore a different approach to the previously mentioned tools, in which one can use the knowledge of absolute Gibbs energies of solvation in the different phases to estimate the partition coefficient. From the Gibbs energies of solvation ($\Delta_{Solv}G$) in two different phases (here generically named *a* and *b*) at constant temperature (*T*), and pressure (*p*), one can readily calculate the corresponding partition coefficient, log $P^{a/b}$, according to the following expression

$$\log P^{a/b} = \frac{\Delta_{solv}G^a - \Delta_{solv}G^b}{2.303RT} \quad (1)$$

Equation 1 requires knowledge of the Gibbs energy of solvation of the solute in both solvents of interest, which can be estimated from molecular simulations of free energies as explained in the next section. Initially, this approach was used to predict the water/CCl_4 partition coefficients of small alcohols[24] and it has previously been applied only to a few specific solvent pairs, e.g., chloroform/water[25] and *n*-octane/water[26] partition coefficients. Recently, this approach has been generalized and applied to the partitioning in solvents of different polarity.[27]

In this study, we aim to explore the prediction of both 1-octanol/water and *n*-hexane/water (log P_{HW}) partition coefficients of different chemicals that are environmentally relevant, namely, substituted alkyl-aromatic molecules, chlorobenzenes, polychlorinated biphenyl (PCB), and polychlorinated diphenyl ether (PCDE) congeners. In Figures 1 and 2, the general structure of PCBs and PCDEs and the atom numbering used in this work are shown. PCBs are hazardous molecules and their environmental impact associated with risk assessment is well documented.[28] PCDEs result from incomplete combustion, e.g., emissions from water incinerators, and have been detected in several environmental samples.[29] Due to their importance, several attempts to predict the partitioning of such chemicals have been undertaken: parameters for the Abraham solvation equation (an empirical relationship) have been recently updated in order to reproduce PCBs log P_{OW}[30]; Van Noort has derived QSPR to predict PCB's log P_{OW} reporting deviations of 0.12 log *P* units[31]; a similar study was presented by Lu et al. where molecular descriptors such as molecular surface area, polarizability and the maximum valency of a carbon atom have been used[32]; very recent QSPR/QSAR models to predict biological activity of PCDEs[33] were proposed; finally, LSERs have also been applied for polycyclic aromatic hydrocarbons and PCBs.[34]

Another class of compounds studied here are halogenated organic compounds, such as chloro- and multichloro-substituted benzenes, which resist degradation, are lipophilic, and

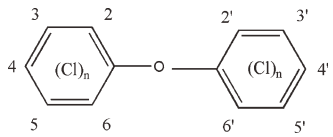

Figure 2. Polychlorinated diphenyl ether (PCDE) structure and atom numbering.

tend to bioaccumulate, being often considered as a hazardous chemical group in toxicology and public health. Some other molecules, involving common functional groups such as alcohols, alkanes and ketones were also included in the test set for validation purposes, since solvation data is widely available for these simpler molecules.

In this article, we test the log P_{OW} and the log P_{HW} predictive capability of molecular simulations, identify successes and failures, and compare its accuracy with that of common methods to estimate log P. The remainder of this paper is organized as follows: in the next section we briefly summarize the computational methods used, and briefly present the force field parameters used in this work. In the Results and Discussion section, we first evaluate the capability of the force field to predict absolute solvation energies, and subsequently discuss partition coefficient predictions. The main conclusions are summarized in the last section.

Simulation Methods and Potential Details

In this work solvation Gibbs free energies have been estimated via molecular simulation and thermodynamic integration (TI), following the protocol presented in our previous articles.[27,35] This methodology has been shown to yield accurate values of solvation free energies in a wide range of environments.[27,36] Using a different approach, partition coefficients can also be directly calculated from the ratio of number densities obtained in a Gibbs ensemble Monte Carlo (GEMC) simulation.[37] In this section, a brief overview of molecular modeling and free energy calculations is provided together with details of the molecular dynamics (MD) protocol.

In classical statistical mechanics, systems are described by their Hamiltonian $\mathcal{H}$, written as a sum of kinetic and potential energy contributions. Potential energy plays a very important role in molecular systems and needs to be accurately determined. It is commonly decomposed into intermolecular energy, arising from the interaction between atoms in different molecules, and intramolecular energy, arising from the interactions between atoms in the same molecule. Intermolecular energy is usually split into an electrostatic component, a polarization component, a dispersion component and a repulsive component. Polarizable models[38] take all four components into account, but are very computationally intensive for the purposes of this study. In this work, we have, thus, employed a nonpolarizable additive force field, which means that the polarizable term is implicitly accounted for by a correct parameterization of the effective point charges in the model. Electrostatic interactions are modeled based on Coulomb's law while the dispersion and repulsive energies are both included into the widely employed Lennard-Jones (LJ) potential model.

The solvation process consists of the transfer of a solute molecule from a well-defined gas (or vacuum) state into solution. For the prediction of solvation Gibbs energies, we have followed the thermodynamic integration approach, where the solvation Gibbs energy can be computed from a thermodynamic cycle in which solute-solvent intermolecular interactions are progressively switched off.[39] One may use a coupling parameter to enable the transition between the two states (a fully interacting state and a noninteracting state where LJ and electrostatic interactions between solute and solvent are zero), and equilibrium averages are used to evaluate derivatives of the free energy with respect to that order parameter. During the decoupling process, the electrostatic interactions were linearly interpolated, as a function of λ, between neighboring states while the LJ interactions were interpolated via soft-core interactions,[40] using a soft-core parameter of 0.5.[41] One then integrates the free energy derivatives along a continuous path connecting the initial and final states to obtain the free energy difference between them. The integration of the LJ Hamiltonian derivatives was carried out by fitting the data to a physically-based approximation to the cavity formation and dispersion interaction terms, and then integrating the curve analytically.[42] For the integration of the electrostatic component, Simpson's rule was used. Here, we simply present the final expression based on which Gibbs energies of solvation are computed

$$\Delta_{Solv}G = \Delta_{Solv}G^{Elec} + \Delta_{Solv}G^{LJ} = -\left[\int_0^1 \left\langle \frac{\partial \mathcal{H}(\lambda)}{\partial \lambda} \right\rangle_\lambda^{Elec} d\lambda + \int_0^1 \left\langle \frac{\partial \mathcal{H}(\lambda)}{\partial \lambda} \right\rangle_\lambda^{LJ} d\lambda \right] \quad (2)$$

Based on the latter, one can compute the Gibbs energy of solvation as a sum of two well-defined contributions (Lennard-Jones and electrostatic interactions). $\mathcal{H}$ is the total Hamiltonian of the system, as a function of the atom coordinates and momenta, and λ is the coupling parameter used to switch off intermolecular interactions. The electrostatic term represents the free energy cost of charging/discharging the solute, while the LJ component is the sum of an unfavorable cavity formation term and a favorable dispersion contribution.

The core of any molecular simulation is the force-field, which represents the interactions between all pairs of atoms in the system. Except for water and chlorine, all atoms (or beads) of the molecules under study were represented by the transferable potential for phase equilibria (TraPPE) force field.[43] In the general framework of this force field, nonbonded interactions are described by pairwise-additive LJ 12-6 potentials and Coulombic interactions of partial charges. We have used a united-atom (UA) description of alkyl CH_x groups[43,44] and an all-atom (AA) description for the polar groups: OH group in 1-octanol[45] and the CO group in ketones and esters.[46] For aromatic rings, we have followed an explicit-hydrogen (EH) description[47] together with a recent parameterization of its partial atomic charges, designed to yield better agreement with experimental Gibbs energies of hydration.[48] This reparameterization slightly underestimates the enthalpies of vaporization relative to the original TraPPE-EH parameters, but correctly reproduces experimental liquid densities. The performance of this model in reproducing vapor–liquid coexistence curves was not assessed. Lorentz-Berthelot combining rules were used to determine the parameters for unlike LJ interactions. All bond lengths were treated as rigid and, with the exception of the fully rigid aromatic rings, all molecular models included harmonic angle bending and torsional degrees of freedom.

Water was represented by the modified extended simple point charge (MSPC/E)[49] model, while aromatic chlorine atoms, due to the unavailability of TraPPE-EH parameters, were represented by the model of Jorgensen et al.[50] The chlorine partial atomic charge was optimized to reproduce the gas-phase dipole moment and later scaled by a constant to reproduce liquid properties.[50] For ketones, improved partial atomic charges assigned to both carbon and oxygen

Table 1. Experimental and Predicted Gibbs Energies of Solvation (kJ/mol) in *n*-Hexane for Different Solutes

Homologous series	Solute	$\Delta_{Solv}\,G^{Exp}$	$\Delta G^{LJ} = \Delta_{Solv}\,G^{Calc}$	Homologous Series Overall AAD (kJ/mol)
1-Alcohols	methanol	−6.2	-3.3_3	3.4
	ethanol	−11.0	-6.6_3	
	1-butanol	−15.8	-12.7_4	
	1-hexanol	−21.5	-18.1_3	
Aromatics	benzene	−16.6	-18.2_3	1.5
	toluene	−20.3	-20.3_4	
	o-xylene	−21.8	-24.2_2	
	m-xylene	−20.9	-24.1_2	
Ketones	acetone	−10.9	-11.4_2	0.6
	diethylketone	n.a.	-17.3_2	
Chlorobenzenes	clorobenzene	−21.5	-22.3_3	1.3
	1,2-dichlorobenzene	n.a.	-26.2_2	
	1,4-dichlorobenzene	−23.8	-25.6_2	
	1,2,3-trichlorobenzene	n.a.	-29.6_2	
	1,3,5-trichlorobenzene	n.a.	-29.9_2	
	hexachlorobenzene	−42.6	-41.2_3	
PCB's	biphenyl	n.a.	-38.0_2	–
	2-PCB	n.a.	-40.9_2	
	4-PCB	n.a.	-41.6_3	
	2,2′-PCB	n.a.	-42.6_3	
	2,5-PCB	n.a.	-44.5_3	
	2,2′,3-PCB	n.a.	-46.3_3	
	2,3,3′,4,4′,5-PCB	n.a.	-59.4_3	
PCDE's	diphenyl ether	n.a.	-38.3_2	–
	2-PCDE	n.a.	-40.1_3	
	4-PCDE	n.a.	-42.0_3	
	2,3′,4-PCDE	n.a.	-49.3_2	
	3,4,4′-PCDE	n.a.	-49.9_3	

The subscripts give the statistical accuracy of the last decimal point shown.

atoms have also been employed, to reproduce correct hydration free energies in the aqueous phase.[27]

Integration of MD equations was conducted with version 4.0.7 of the GROMACS software[51] using the leap-frog Verlet integration algorithm[52] with a time step of 2 fs. Simulations were performed using periodic boundary conditions in all directions. Covalent bonds involving hydrogen atoms were constrained using the LINCS algorithm[53] while the water geometry was fixed with the SETTLE algorithm[54]. For efficiency reasons (see[55] for details), the reaction-field method,[56] which approximates the medium beyond a cut-off distance of 1 nm by a dielectric continuum of uniform permittivity, was used to handle long-range electrostatics. The dielectric constant was adjusted to be the permittivity of each solvent. The remaining cut-off radii were 1 nm for the short-range neighbor list and a 0.8–0.9 nm switched cut-off for the LJ interactions. We have also applied long-range dispersion corrections for energy and pressure.[57] Solvated systems consisted of one solute molecule and 500 water, 150 1-octanol or 156 hexane molecules at 298 K and 1 bar.

Langevin stochastic dynamics[58] was used to control the temperature, with a frictional constant of 1 ps^{-1}, while for constant pressure runs the Berendsen barostat,[59] with a time constant of 0.5 ps, and an isothermal compressibility of 4.5×10^{-5} bar^{-1}, was used to enforce pressure coupling. For each simulation, we first run an energy minimization (the Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm[60] during 5,000 steps followed by a steepest descent minimization during 500 steps were used) and afterward a constant volume equilibration (100 ps), a constant pressure equilibration (500 ps), sufficient to obtain complete equilibration of the box volume, and finally a 5 ns *NpT* production stage.

Results and Discussion

Gibbs energies of solvation

In the first part of this work, we have estimated the solvation Gibbs energies of the different solutes in *n*-hexane, water and 1-octanol using the previously mentioned thermodynamic integration methodology. Results obtained are presented in Tables 1 through 3, respectively, for the three solvents. For the case of polar solvents, both LJ and electrostatic contributions to the total solvation energy are given. Experimental data[61] are also included, when available. Unfortunately, available experimental data do not report their associated uncertainty. In general, uncertainty estimates for some common compounds have been suggested to be around 0.2 kcal/mol or higher.[62,63]

Since *n*-hexane is a nonpolar molecule, the Coulombic term in Eq. 2 becomes zero. This means that simulations in *n*-hexane are particularly useful because they enable us to evaluate solely the influence of the LJ intermolecular interactions, separated from the effect of the point charge parameterization. Experimental data for the monofunctional molecules are widely available and one can compare simulation predictions directly against such data. This is the main reason why we need to include simpler molecules in our test set. In general, experimental solvation energies are well predicted by the TraPPE force-field. The only exception is the

Table 2. Experimental and LJ and Coulombic Contributions to the Predicted Gibbs Energies of Hydration (kJ/mol) for the Different Compounds

Homologous series	Solute	$\Delta_{Hyd}G^{Exp}$	ΔG^{LJ}	ΔG^{Coul}	$\Delta_{Hyd}G^{Calc}$	Homologous Series Overall AAD (kJ/mol)
1-Alcohols	methanol	−21.2	11.6_3	-30.43_6	-19.0_3	2.2
	ethanol	−21.1	8.0_2	-28.33_5	-20.1_2	
	1-butanol	−19.8	10.6_5	-27.96_5	-17.3_5	
	1-hexanol	−18.0	13.2_4	-28.14_5	-14.9_4	
Aromatics	benzene	−3.6	5.5_2	-9.13_3	-3.6_2	0.4
	toluene	−3.7	5.9_2	-10.15_3	-4.2_2	
	ethylbenzene	−3.3	7.1_2	-10.32_4	-3.3_2	
	o-xylene	−3.8	6.6_2	-10.8_3	-4.2_2	
	m-xylene	−3.5	7.1_2	-11.2_2	-4.1_2	
	p-xylene	−3.4	6.6_2	-10.9_3	-4.3_2	
Ketones	acetone	−15.9	6.4_4	-22.40_5	-16.0_4	0.05
	diethylketone	n.a.	6.4_4	-23.66_6	-17.2_4	–
Chlorobenzenes	clorobenzene	−5.2	6.0_3	-6.33_3	-0.3_3	2.2
	1,2-dichlorobenzene	−6.3	4.1_3	-8.65_3	-4.6_3	
	1,4-dichlorobenzene	−5.8	5.1_3	-8.75_3	-3.7_3	
	1,2,3-trichlorobenzene	−7.4	5.0_3	-12.43_4	-7.5_3	
	1,3,5-trichlorobenzene	n.a.	5.4_3	-12.61_4	-7.2_3	
	hexachlorobenzene	n.a.	2.4_4	-31.32_8	-28.9_4	
PCB's	biphenyl	n.a.	3.4_3	-17.84_4	-14.4_3	–
	2-PCB	n.a.	4.2_3	-17.16_4	-13.0_3	
	4-PCB	n.a.	5.3_3	-16.36_4	-11.1_3	
	2,2′-PCB	n.a.	5.9_3	-13.88_5	-8.0_3	
	2,5′-PCB	n.a.	4.4_3	-17.53_6	-13.2_3	
	2,2′,3-PCB	n.a.	3.3_4	-14.77_5	-11.4_4	
	2,3,3′,4,4′,5-PCB	n.a.	2.0_3	-30.74_7	-28.7_3	
PCDE	diphenyl ether	n.a.	7.5_3	-28.00_8	-20.5_3	–
	2-PCDE	n.a.	7.9_3	-28.92_7	-21.1_3	
	4-PCDE	n.a.	7.9_3	-26.30_7	-18.4_3	
	2,3′,4-PCDE	n.a.	7.3_4	-27.30_7	-20.0_4	
	3,4,4′-PCDE	n.a.	5.9_3	-27.03_8	-21.1_3	

The subscripts give the statistical accuracy of the last decimal point shown.

Table 3. Experimental and LJ and Coulombic Contributions to the Predicted Gibbs Energies of Solvation (kJ/mol) in 1-Octanol for the Different Compounds

Homologous series	Solute	$\Delta_{Solv}G^{Exp}$	ΔG^{LJ}	ΔG^{Coul}	$\Delta_{Solv}G^{Calc}$	Homologous Series Overall AAD (kJ/mol)
1-Alcohols	methanol	−16.3	-2.0_3	-13.53_5	-15.6_3	1.0
	ethanol	−18.4	-4.4_3	-14.11_5	-18.4_3	
	1-butanol	−24.5	-9.9_4	-11.26_3	-21.2_4	
	1-hexanol	−29.5	-17.7_4	-10.00_3	-27.7_4	
Aromatics	benzene	−15.9	-16.1_3	-0.69_2	-16.8_3	0.7
	toluene	−18.9	-19.2_4	-0.67_3	-19.9_4	
	ethylbenzene	−21.2	-20.6_3	-1.11_6	-21.7_3	
	o-xylene	−22.2	-21.5_4	-1.19_5	-22.7_4	
	m-xylene	−21.6	-20.6_3	-1.16_5	-21.8_3	
	p-xylene	−21.6	-21.6_3	-1.33_5	-22.9_3	
Ketones	acetone	−13.2	-9.9_2	-2.29_5	-12.2_2	1.0
	diethylketone	n.a.	-14.5_3	-2.33_4	-16.8_3	–
Chlorobenzenes	clorobenzene	n.a.	-20.4_5	-1.65_4	-22.1_5	
	1,2-dichlorobenzene	−24.9	-25.4_5	-2.74_5	-28.1_5	
	1,4-dichlorobenzene	−25.4	-25.1_3	-1.47_3	-26.6_3	
	1,2,3-trichlorobenzene	−29.6	-28.4_5	-3.57_5	-32.0_5	
	1,3,5-trichlorobenzene	−27.7	-28.0_5	-1.32_4	-29.3_5	
	hexachlorobenzene	−42.0	-39.8_7	-18.66_7	-58.5_7	
PCB's	biphenyl	−35.1	-37.0_5	-2.98_3	-40.0_5	4.9
	2-PCB	n.a.	-39.7_4	-2.90_3	-42.6_4	
	4-PCB	n.a.	-39.9_5	-2.74_3	-42.7_5	
	2,2′-PCB	n.a.	-41.5_5	-2.89_4	-44.3_5	
	2,5-PCB	n.a.	-43.6_6	-3.04_3	-46.6_6	
	2,2′,3-PCB	n.a.	-46.2_6	-5.39_3	-51.6_6	
	2,3,3′,4,4′,5-PCB	n.a.	-60.5_7	-13.85_5	-74.4_7	
PCDE's	diphenyl ether	n.a.	-36.6_5	-5.26_4	-41.9_5	–
	2-PCDE	n.a.	-38.7_5	-5.43_5	-44.1_5	
	4-PCDE	n.a.	-40.2_5	-5.22_3	-45.5_5	
	2,3′,4-PCDE	n.a.	-46.3_6	-3.81_4	-50.1_6	
	3,4,4′-PCDE	n.a.	-48.5_5	-6.29_5	-54.8_5	

The subscripts give the statistical accuracy of the last decimal point shown.

case of alcohols where slightly larger deviations (AAD around 3 kJ/mol) are found. However, this may be due to a large uncertainty in the experimental values, as already discussed in a previous work.[48] The main argument in support of this hypothesis is that accurate log P_{OW} and log P_{HW} data are predicted from the simulations, despite the discrepancies between experimental and simulated hydration free energies. This effect was observed previously,[27] and is discussed in more detail later (see Partition Coefficients section).

For the more complex molecules, experimental data are scarce, but in the case of chlorinated benzenes the experimental trend is very well reproduced, particularly for the case of hexachlorobenzene which has the most favorable solvation energy of the series. In short, the results from Table 1 enable us to conclude that the LJ term of the TraPPE force field accurate represents the nonpolar (dispersion and repulsion) contributions to the total potential energy.

We turn our attention next to the simulations in the water phase, a highly polar solvent. Results obtained are also in very good agreement with experimental data, indicating that the description of the polar groups (namely by their assigned atomic partial charges) is generally correct. In the case of the chlorinated benzene family it seems that simulation predictions overestimate the hydration energy values, although the obtained average absolute deviation of 2.2 kJ/mol can be considered reasonably accurate if one compares it with standard deviations in other hydration studies from molecular simulation.[57,64,65] The accuracy of these values is reflected in the good log P predictions for this class of compounds (see Partition Coefficients section). However, we should note that the parameterization of the aromatic chlorine atom was done only considering the chlorobenzene molecule[50] and did not explicitly consider multisubstituted halogenated aromatics. This effect manifests itself in the very high value obtained for the electrostatic term in hexachlorobenzene.

In this work, we assume that the chlorine charge value remains constant, independent of the degree of aromatic substitution, which is a rough approximation. At this point, it becomes important to explore and discuss further how force fields are developed and how the transferability of parameters between different molecules may become problematic. Two key factors are usually involved in the successful use of force field parameters: accuracy and transferability. If one is unable to achieve an accurate fit to the training data, it is unlikely that the force field will perform well for more complex systems. However, even if a parameter set is successful in accurately reproducing the behavior of the model systems, testing transferability to the larger systems and properties of interest is of critical importance.[66] During the parameterization procedure of a given force field, LJ sites and intramolecular bonded energetic parameters can be reasonably (and automatically) assigned through atom-typing and the use of hierarchy well-defined rules.[67] The principal problem is the correct assignment of partial atomic charges, particularly for new or arbitrary molecules. Such charges can be obtained, e.g., from gas-phase *ab initio* potentials, quantum mechanics calculations or can be fitted to reproduce one or more liquid-phase experimental properties. In the case of the Jorgensen model for chlorine aromatic atoms used here, following the OPLS philosophy[68] the charge parameterization strategy is based on a balance between the reproduction of gas-phase dipole moments, pure liquid properties and aqueous behavior. This strategy has a drawback, because if we deal with an organic molecule in an environment with some degree of polarization, the charge distribution selected will result in an absolute free energy that is smaller than expected.[69]

The solution to overcome this problem and the choice of suitable parameters to work on different environments is far from trivial. One approach is to include explicitly polarization in the force field model. Another solution is to assign different partial charges to the solute depending on the environment. In this case, solvation data may also be included in the parameterization of partial atomic charges.[48,70] However, a generic application of this expensive parameterization procedure to a large data set (to decrease the risk of over-fitting to a single experiment data point with an unknown uncertainty) using free energy techniques is still today very demanding and time-consuming.

The transferability of force field parameters becomes notoriously problematic if one tries to apply the same atom type values in a broad range of molecules. For example, chlorobenzene, *o*-, *m*-, and *p*-dichlorobenzene conceivably require different partial charges on the chlorine atom[69]; however, it is possible to choose a single-partial charge value that will perform reasonably well regardless of the presence and position of the second substitution. The large discrepancies observed with experiment for the solvation of hexachlorobenzene in 1-octanol (Table 3) may suggest that separate atomic parameters should be considered. This observation is consistent with the parameterization of fluorobenzenes with the OPLS-AA force field,[50] where fluorobenzene, difluorobenzene and hexafluorobenzene have progressively smaller partial charges on the fluorine atoms to account for the natural change in charge distribution from increased substitution.[69] However, such a detailed parameterization is outside of the scope of this work.

In terms of comparison, literature reported values[69] for the prediction of relative hydration Gibbs energy differences between benzene and chlorobenzene are 1.2_4 kJ/mol (the subscripts give the statistical accuracy of the last decimal point) while the experimental value is -0.50 kJ/mol. It is worthwhile to notice that this prediction reflects only the value of mutating from benzene to chlorobenzenes, which is natural easier to compute. When mutating between clorobenzene and 1,2–, 1,3– and 1,4–dichlorobenzene predicted values from Ref. 69 are -3.7, -0.04 and -1.3 kJ/mol, while experimental values from the same reference are -1.5, -0.1 and 0.0 kJ/mol, respectively.

Finally, in Table 3 we report results of MD simulations in the 1-octanol phase. Once again, for the monofunctional solutes excellent predictions are obtained. Our predictions are also close to the predicted solvation Energies of Chen and Siepmann,[3] who used GEMC simulations and the TraPPE force field. In the case of the chlorinated benzenes, results predicted for the different compounds are in good agreement with experimental data, except in the case of hexachlorobenzene. This fact may again indicate that partial atomic charges assigned to chlorine atoms connected to the aromatic ring are perhaps too large in magnitude to reproduce solvation energies. In fact, this molecule could be used to obtain the optimal charges assigned to chlorine substituents that exactly reproduce solvation energies in 1-octanol.

Overall, predicted values can be considered reasonably accurate given that uncertainties in the experimental data are about ± 1.2 kJ/mol.[71] The remaining differences between experimental values and MD calculations can be attributed

Table 4. Experimental[10,17] and Predicted *n*-Hexane/Water Partition Coefficients Data (log P_{HW}) from MD Simulations

	log P_{HW}	
Solute	Experimental	Simulation
methanol	−2.80	−2.75
ethanol	−2.26	−2.37
1-butanol	−0.78	−0.81
1-hexanol	0.45	0.56
benzene	2.06	2.56
toluene	2.75	2.83
o-xylene	3.12	3.52
m-xylene	3.04	3.51
acetone	−0.92	−0.79
diethylketone	n.a.	0.02
clorobenzene	3.00	3.85
1,2-dichlorobenzene	3.60	3.78
1,4-dichlorobenzene	3.81	3.85
1,2,3-trichlorobenzene	4.16	3.88
1,3,5-trichlorobenzene	4.53	3.98
hexachlorobenzene	5.69	2.15
biphenyl	n.a.	4.13
2-PCB	n.a.	4.90
4-PCB	n.a.	5.35
2,2′-PCB	n.a.	6.06
2,5-PCB	n.a.	5.50
2,2′,3-PCB	n.a.	6.10
2,3,3′,4,4′,5-PCB	n.a.	5.37
diphenyl ether	n.a.	3.12
2-PCDE	n.a.	3.84
4-PCDE	n.a.	4.14
2,3′,4-PCDE	n.a.	5.13
3,4,4′-PCDE	n.a.	5.05
AAD		0.48

Table 5. Experimental[5,19] and Predicted 1-Octanol/Water Partition Coefficients Data (log P_{OW}) from MD Simulations

	log P_{OW}	
Solute	Experimental	Simulation
methanol	−0.74	−0.59
ethanol	−0.30	−0.30
1-butanol	0.84	0.68
1-hexanol	2.03	2.24
methane	1.09	1.32
ethane	1.81	1.69
propane	2.36	2.27
n-butane	2.89	2.64
n-pentane	3.39	3.52
n-hexane	3.90	4.11
n-heptane	4.66	4.62
n-octane	5.15	5.13
benzene	2.13	2.31
toluene	2.73	2.75
ethylbenzene	3.15	3.22
o-xylene	3.12	3.24
m-xylene	3.20	3.10
p-xylene	3.15	3.26
acetone	−0.24	−0.66
diethylketone	0.82	−0.07
chlorobenzene	2.84	3.82
1,2-dichlorobenzene	3.38	4.13
1,4-dichlorobenzene	3.45	4.01
1,2,3-trichlorobenzene	4.05	4.30
1,3,5-trichlorobenzene	4.15	3.88
hexachlorobenzene	5.44	5.19
biphenyl	4.09	4.48
2-PCB	4.38	5.19
4-PCB	4.63	5.36
2,2′-PCB	4.90	6.37
2,5-PCB	5.00	5.86
2,2′,3-PCB	5.60	7.04
2,3,3′,4,4′,5-PCB	7.44	8.00
diphenyl ether	n.a.	3.75
2-PCDE	4.45	4.05
4-PCDE	4.79	4.76
2,3′,4-PCDE	5.65	5.26
3,4,4′-PCDE	5.88	5.91
AAD		0.41

primarily to a deficiency in the charge distributions of the model. Further studies are needed to clarify this issue.

Partition coefficients

The *n*-hexane/water and 1-octanol/water partition coefficients at 298 K and 1 bar for the full test set can be readily estimated from Eq. 1 using the Gibbs energies of solvation previously calculated from the MD simulations. Results are shown for the different solutes in Tables 4and 5 and represented in Figures 3 and 4, for log P_{HW} and log P_{OW}, respectively. The available experimental data are also included in the tables. However, it is important to note that available log P_{OW} data may differ significantly depending on the experimental method employed. For instance, measured log P_{OW} values for a molecule as simple as benzene range from 1.56 to 2.34 as discussed in Garst and Wilson,[72] whereas reported data for hexachlorobenzene range from 4.13 to 7.42.[73] In Tables 4 and 5 we have included the suggested experimental values.

Overall average absolute deviations (AAD) of 0.48 log P_{HW} and 0.41 log P_{OW} units were obtained when compared to experimental data. From Figure 3 one can verify that one of the data points exhibits a large deviation from experimental data. This point refers to hexachlorobenzene. By removing this point from the test set, the AAD to experiments decreases to 0.27 log P_{HW} units. The poor description of this log P_{HW} is probably due to a large deviation of the solvation free energy in water (experimental value is not available) while the *n*-hexane phase is likely to be well described, as already discussed in the section "Gibbs Energies of Solvation".

Deviations presented here are in general higher than the ones obtained from statistical methods (ca. 0.2 log *P* units), but it is worthwhile to notice that our methodology is purely predictive (without need for experimental input), and can be

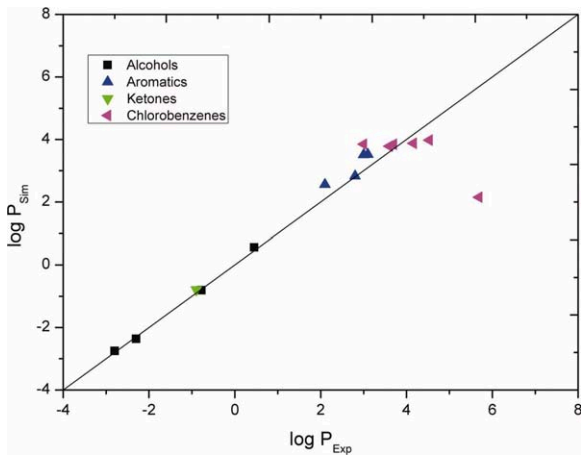

Figure 3. Comparison of experimental data vs. simulation predictions for log *P* (*n*-hexane/water).

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

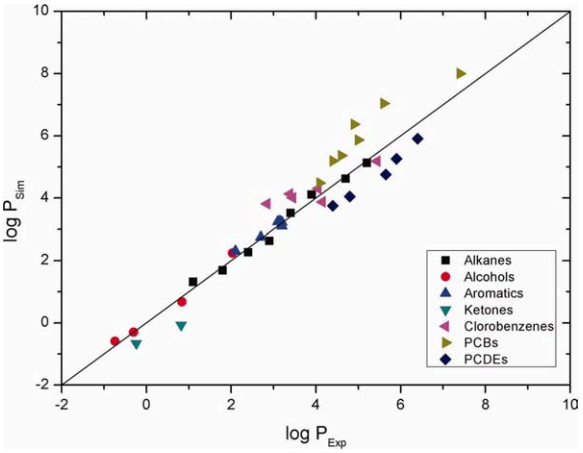

Figure 4. Comparison of experimental data vs. simulation predictions for log *P* (1-octanol/water).

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

applied to any pair of solute/solvent systems without the knowledge of a wide range of solute properties used for statistical analysis in QSPR methods. Apart from QSPR studies, in terms of accuracy comparison, and although applied to larger solute databases, deviations of 0.46 and 0.43 log *P* units for 1-octanol/water and *n*-hexane/water partition, respectively, can be found in the literature using an empirical thermodynamic approach.[19] Multiple regression techniques applied on descriptors based on molecular properties and molecular partial charges can produce deviations of 0.65 log *P* units for 1-octanol/water partition coefficient.[74] For the same property, a combination of a group-contribution approach together with several correction factors produces deviations of 0.4 log *P* units.[75] Neural network techniques applied to the prediction of log P from nonlinear regression of structure analysis using 10 to 18 descriptors produced deviations ranging from 0.3 to 0.4 log *P* units.[76] Finally, a generic test of quantum mechanical and QSPR methodologies produced average deviations of 0.6 and a maximum of 2.15 log *P* units.[77] This comparison suggests that molecular simulation methods are competitive in terms of accuracy with other techniques that are more commonly employed to predict partition coefficients.

Conclusions

The focus of this work has been the prediction of the log P_{HW} and log P_{OW} for several environmentally important solutes, through the calculation of the absolute Gibbs energies of solvation using MD simulations and thermodynamic integration. We have tested the Gibbs energy predictive capability of the TraPPE force field, used to account for the intra- and intermolecular interactions, identifying success and failure cases, and pointing out particular classes of compounds that could benefit from a refinement of the force field parameters. Some of the failures presented in this article, namely for chlorobenzenes, may be associated with using OPLS-AA LJ parameters and partial charges for the chlorine atom together with TraPPE parameters for the aromatic ring (mixing parameters from different force fields is always questionable). Thus, it remains an open question whether the newly developed TraPPE-EH parameters for chlorinated aromatics would suffer from similar shortcomings. The accuracy of the TraPPE force field for the prediction of Gibbs energies and partition data has not yet been thoroughly evaluated, and the calculations presented here represent an important step in this direction.

In general, our predictive results for log P_{HW} and log P_{OW}, which did not include correlation of solvation or partitioning experimental data, are in satisfactory agreement with experimental data. The accuracy of our predictions is comparable to the popular statistically-based methods used to generate such partition data. Furthermore, the approach followed here has the additional advantage of being easily extended to any pair of binary solvent systems.

From the results in the inert (*n*-hexane) phase, one concludes that nonpolar parameters of the force field used here can correctly describe the relevant microscopic thermodynamic behavior. In the case of the poly-chlorobenzenes the description of the solvation free energy in polar solvents can be improved by assigning different partial atomic charges depending on the degree of substitution. With the increase in computational power and with the availability of more experimental data it will become possible to directly include solvation energy data in the parameterization of a force field.

In view of the recent advances in available computing power, molecular simulation is emerging as a powerful computational tool, allowing for fully predictive results to be achieved in a reasonable amount of time. However, a prediction of the partition coefficient for a single solute can take around 24 h using a state-of-the-art processor. In this respect, we cannot use such larger databases as many times as when statistical methods are used to correlate the corresponding data. However, we believe that the test set presented in this work is large enough to cover a wide variety of cases and molecules important for environmental applications.

Acknowledgments

NMG acknowledges his FCT PhD scholarship SFRH/BD/47822/2007 and Petroleum Institute for financial support through a visiting PhD scholarship.

Literature Cited

1. Leo A, Hansch C, Elkins D. Partition coefficients and their uses. *Chem Rev*. 1971;71(6):525–616.
2. Debolt SE, Kollman PA. Investigation of structure, dynamics and solvation in 1-octanol and its water-saturated solution: molecular dynamics and free-energy perturbation studies. *J Am Chem Soc*. 1995;117(19):5316–5340.
3. Chen B, Siepmann JI. Microscopic structure and solvation in dry and wet octanol. *J Phys Chem B*. 2006;110(8):3555–3563.
4. Hansch C, Leo A, Hoekman D. *Exploring QSAR: Hydrophobic, Electronic and Steric Constants*. Washington, DC: American Chemical Society; 1995.
5. Sangster J. *Octanol-Water Partitioning Coefficients: Fundamentals and Physical Chemistry*. Chichester, U.K: John Wiley & Sons; 1997.
6. Pinho SP, Macedo EA. *Solubility in Food, Pharmaceuticals, and Cosmetic Industries*. in: *Developments and Applications in Solubility*. Letcher TM, ed. Cambridge: Royal Society of Chemistry; 2003:309–326.
7. Yalkowsky SH. *Solubility and Solubilization in Aqueous Media*. Oxford: Oxford University; 1999.
8. Chiou CT, Freed VH, Schmedding DW, Kohnert RL. Partition coefficient and bioaccumulation of selected organic chemicals. *Environ Sci Techn*. 1977;11(5):475–478.
9. Mackay D. Correlation of bioconcentration factors. *Environ. Sci. Techn*. 1982;16(5):274–278.
10. Schulte J, Durr J, Ritter S, Hauthal WG, Quittzsch K, Maurer G. Partition coefficients for environmentally important, multifunctional

organic compounds in hexane + water. *J Chem Eng Data.* 1998;43:69–73.
11. Bergstrom CAS, Norinder U, Luthman K, Artursson P. Experimental and computational screening models for prediction of aqueous drug solubility. *Pharm Res.* 2002;19(2):182–188.
12. Glomme A, Marz J, Dressman JB. Comparison of a miniaturized shake-flask solubility method with automated potentiometric acid/base titrations and calculated solubilities. *J Pharm Sci.* 2005;94(1):1–16.
13. Hansch C, Fujita T. ρ-π-σ analysis. A Method for the correlation of biological activity and chemical structure. *J Am Chem Soc.* 1964;86(8):1616–1626.
14. Leo AJ. Calculating log P(Oct) from structures. *Chem Rev.* 1993;93(4):1281–1306.
15. Viswanadhan VN, Ghose AK, Singh UC, Wendoloski JJ. Prediction of solvation free energies of small organic molecules: Additive-constitutive models based on molecular fingerprints and atomic constants. *J Chem Inf Comput Sci.* 1999;39(2):405–412.
16. Johnson S. The trouble with QSAR (or how i learned to stop worrying and embrace fallacy). *J Chem Inf Model.* 2008;48:25–26.
17. Goss K, Schwarzenbach RP. Linear free energy relationships used to evaluate equilibrium partitioning of organic compounds. *Environ Sci Techn.* 2001;35(1):1–9.
18. Nguyen TH, Goss K, Ball WP. Polyparameter linear free energy relationships for estimating the equilibrium partition of organic compounds between water and the natural organic matter in soils and sediments. *Environ Sci Techn.* 2005;39(4):913–924.
19. Ruelle P. The n-octanol and n-hexane/water partition coefficient of environmentally relevant chemicals predicted from the mobile order and disorder (MOD) thermodynamics. *Chemosphere.* 2000;40(5):457–512.
20. Duffy EM, Jorgensen WL. Prediction of properties from simulations: Free energies of solvation in hexadecane, octanol, and water. *J Am Chem Soc.* 2000;122(12):2878–2888.
21. Wang S, Sandler SI, Chen C-C. Refinement of COSMO-SAC and the applications. *Ind Eng Chem Res.* 2007;46(22):7275–7288.
22. Cramer CJ, Truhlar DG. *SMx continuum models for condensed phases.* In: Maroulis G, Simos TE, eds. *Trends and Perspectives in Modern Computational Science; Lecture Series on Computer and Computational Sciences.* Leiden: Brill/VSP; 2006:112–140.
23. Best SA, Merz KM, Reynolds CH. Free energy perturbation study of octanol/water partition coefficients: Comparison with continuum GB/SA calculations. *J Phys Chem B.* 1999;103(4):714–726.
24. Essex JW, Reynolds CA, Richards WG. Theoretical determination of partition coefficients. *J Am Chem Soc.* 1992;114(10):3634–3639.
25. Eksterowicz JE, Miller JL, Kollman PA. Calculation of chloroform/water partition coefficients for the N-methylated nucleic acid bases. *J Phys Chem B.* 1997;101(50):10971–10975.
26. Michel J, Orsi M, Essex JW. Prediction of partition coefficients by multiscale hybrid atomic-level/coarse-grain simulations. *J Phys Chem B.* 2008;112(3):657–660.
27. Garrido NM, Jorge M, Queimada AJ, Macedo EA, Economou IG. Using molecular simulation to predict solvation and partition coefficients in solvents of different polarity. *Phys Chem Chem Phys.* 2011;13:9155–9164.
28. Safe SH. Polychlorinated biphenyls (PCBs): Environmental impact, biochemical and toxic responses, and implications for risk assessment. *Crit Rev Toxic.* 1994;24(2):87–149.
29. Lyytikainen M, Hirva P, Minkkinen P, Hamalainen H, Rantalainen A, Mikkelson P, Paasivirta J, Kukkonen JVK. Bioavailability of sediment-associated pcdd/fs and pcdes: relative importance of contaminant and sediment characteristics and biological factors. *Environ Sci Technol.* 2003;37(17):3926–3984.
30. van Noort PCM, Haftka JH, Parsons JR. Updated abraham solvation parameters for polychlorinated biphenyls. *Environ Sci Technol.* 2010;44:7037–7042.
31. van Noort PCM. Estimation of amorphous organic carbon/water partition coefficients, subcooled liquid aqueous solubilities, and n-octanol/water partition coefficients of nonpolar chlorinated aromatic compounds from chlorine fragment constants. *Chemosphere.* 2009;74:1024–1030.
32. Lu W, Chen Y, Liu M, Chen X, Hu Z. QSPR prediction of n-octanol/water partition coefficient for polychlorinated biphenyls. *Chemosphere.* 2007;69(3):469–478.
33. Hui-Ying X, Jian-Wei Z, Gui-Xiang H, Wei W. QSPR/QSAR models for prediction of the physico-chemical properties and biological activity of polychlorinated diphenyl ethers (PCDEs). *Chemosphere.* 2010;80:665–670.
34. Kamlet MJ, Doherty RM, Abraham MH, Marcus Y, Taft RW. Linear solvation energy relationships 46: an improved equation for correlation and prediction of octanol water partition coefficients of organic nonelectrolytes (including strong hydrogen-bond donor solutes). *J Phys Chem.* 1988;92(18):5244–5255.
35. Garrido NM, Queimada AJ, Jorge M, Macedo EA, Economou IG. 1-Octanol/water partition coefficients of n-alkanes from molecular simulations of absolute solvation free energies. *J Chem Theory Comput.* 2009;5(9):2436–2446.
36. Geerke DP, van Gunsteren WF. Force field evaluation for biomolecular simulation: Free enthalpies of solvation of polar and apolar compounds in various solvents. *Chemphyschem.* 2006;7(3):671–678.
37. Chen B, Siepmann JI. Partitioning of alkane and alcohol solutes between water and (Dry or wet) 1-Octanol. *J Am Chem Soc.* 2000;122(27):6464–6467.
38. Rick SW, Stuart SJ. Potentials and algorithms for incorporating polarizability in computer simulations. *Rev Comput Chem.* 2002;18:89–146.
39. Leach A. *Molecular Modeling: Principles and Applications.* Prentice-Hall; 2001.
40. Beuler TM R, van Schaik RC, Gerber PR, van Gunsteren WF. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem Phys Lett.* 1994;222:529–539.
41. Garrido NM, Jorge M, Queimada AJ, Economou IG, Macedo EA. Molecular simulation of the hydration Gibbs energy of barbiturates. *Fluid Phase Equilib.* 2010;289:148–155.
42. Jorge M, Garrido NM, Queimada AJ, Economou IG, Macedo EA. The effect of integration method on the accuracy and computational efficiency of free energy calculations using thermodynamic integration. *J Chem Theory Comput.* 2010;6(4):1018–1027.
43. Martin MG, Siepmann JI. Transferable potentials for phase equilibria. 1. United-atom description of n-alkanes. *J Phys Chem B.* 1998;102(14):2569–2577.
44. Martin MG, Siepmann JI. Novel configurational-bias Monte Carlo method for branched molecules. Transferable potentials for phase equilibria. 2. United-atom description of branched alkanes. *J Phys Chem B.* 1999;103(21):4508–4517.
45. Chen B, Potoff JJ, Siepmann JI. Monte Carlo calculations for alcohols and their mixtures with alkanes. Transferable potentials for phase equilibria. 5. United-atom description of primary, secondary, and tertiary alcohols. *J Phys Chem B.* 2001;105(15):3093–3104.
46. Stubbs JM, Potoff JJ, Siepmann JI. Transferable potentials for phase equilibria. 6. United-atom description for ethers, glycols, ketones, and aldehydes. *J Phys Chem B.* 2004;108(45):17596–17605.
47. Rai N, Siepmann JI. Transferable potentials for phase equilibria. 9. explicit hydrogen description of benzene and five-membered and six-membered heterocyclic aromatic compounds. *J Phys Chem B.* 2007;111:10790–10799.
48. Garrido NM, Jorge M, Queimada AJ, Economou IG, Macedo EA. Predicting hydration gibbs energies of alkyl-aromatics using molecular simulation: a comparison of current force fields and the development of a new parameter set for accurate solvation data. *Submitted for publication.* 2011.
49. Boulougouris GC, Economou IG, Theodorou DN. Engineering a molecular model for water phase equilibrium over a wide temperature range. *J Phys Chem B.* 1998;102(6):1029–1035.
50. Jorgensen WL, Ulmschneider JP, Tirado-Rives J. Free energies of hydration from a generalized born model and an all-atom force field. *J Phys Chem. B.* 2004;108:16264–16270.
51. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem. Theory Comput.* 2008;4(3):435–447.
52. van Gunsteren W, Berendsen H. A leap-frog algorithm for stochastic dynamics. *Mol Sim.* 1988;1(3):173–185.
53. Hess B, Bekker H, Berendsen HJC, Fraaije J. LINCS: A linear constraint solver for molecular simulations. *J Comp Chem.* 1997;18(12):1463–1472.
54. Miyamoto S, Kollman PA. Settle - an analytical version of the shake and rattle agorithm for rigid water molecules. *J Comp Chem.* 1992;13(8):952–962.
55. Garrido NM, Jorge M, Queimada AJ, Economou IG, Macedo EA. Hydration of substituted barbiturates by molecular simulation: A free energy approach. *Fluid Phase Equilib.* 2010;289:148–155.
56. Lee FS, Warshel A. A local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations. *J Chem Phys.* 1992;97(5):3100–3107.

57. Shirts MR, Pitera JW, Swope WC, Pande VS. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *J Chem Phys*. 2003;119(11):5740–5761.
58. Van Gunsteren WF, Berendsen HJC. Algorithms for brownian dynamics. *Mol Phys*. 1982;45(3):637–647.
59. Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys*. 1984;81(8):3684–3690.
60. Liu DC, Nocedal J. On the limited memory BFGS method for large-scale optimization. *Math Programm*. 1989;45(3):503–528.
61. Katritzky AR, Oliferenko AA, Oliferenko PV, Petrukhin R, Tatham DB, Maran U, Lomaka A, Acree Jr AE. A general treatment of solubility. 1. The QSPR correlation of solvation free energies of single solutes in series of solvents. *J Chem Inf Comput Sci*. 2003;43:1794–1805.
62. Nicholls A, Mobley DL, Guthrie JP, Chodera JD, Bayly CI, Cooper MD, Pande VS. Predicting small-molecule solvation free energies: an informal blind test for computational chemistry. *J Med Chem*. 2008;51:769–779.
63. Abraham MH, Whiting GS, Fuchs R, Chambers EJ. Thermodynamics of solute transfer from water to hexadecane. *J Chem Soc Perkin Trans*. 1990;2:291–300.
64. Shirts MR, Mobley DL, Chodera JD. Alchemical free energy calculations: ready for prime time? *Ann Rep Comput Chem*. 2007;3(4):41–59.
65. Mobley DL, Bayly CI, Cooper MD, Shirts MR, Dill KA. Small molecule hydration free energies in explicit solvent: An extensive test of fixed-charge atomistic simulations. *J Chem Theory Comput*. 2009;5(2):350–358.
66. Okur A, Strockbine B, Hornak V, Simmerling C. Using pc clusters to evaluate the transferability of molecular mechanics force fields for proteins. *J Comput Chem*. 2003;24:21–31.
67. Jorgensen WL, Tirado-Rives J. Molecular modeling of organic and biomolecular systems using boss and mcpro. *J Comput Chem*. 2005;26(16):1689–1700.
68. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc*. 1996;118(45):11225–11236.
69. Price DJ, Brooks III CL. Detailed considerations for a balanced and broadly applicable force field: a study of substituted benzenes modeled with OPLS-AA. *J Comput Chem*. 2005;26:1529–1541.
70. Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J Comp Chem*. 2004;25(13):1656–1676.
71. Dallas AJ, Carr PW. A Thermodynamic and solvatochromic investigation of the effect of water on the phase transfer properties of Ocatn-1-ol. *J Chem Soc Perkin Trans 2*. 1992(12):2155–2161.
72. Garst JE, Wilson WC. Accurate, wide-range, automated, high-performance liquid chromatographic method for the estimation of octanol/water partition coefficients I: Effect of chromatographic conditions and procedure variables on accuracy and reproducibility of the method. *J Pharm Sci*. 1984;73(11):1616–1623.
73. Sabljic A. Chemical topology and ecotoxicology. *Sci Total Environ*. 1991;109/110:197–220.
74. Xing L, Glen RC. Novel methods for the prediction of log P, pKa, and log D. *J Chem Inf Comput Sci*. 2002;42:796–805.
75. Klopman G, Li J-Y, Wang S, Dimayuga M. Computer automated log p calculations based on an extended group contribution approach. *J Chem Inf Comput Sci*. 1994;34:752–781.
76. Duprat AF, Huynh T, Dreyfus G. Toward a principled methodology for neural network design and performance evaluation in QSAR. Application to the prediction of logp. *J Chem Inf Comput Sci*. 1998;38:586–594.
77. Beck B, Breindl A, Clark T. QM/NN QSPR models with error estimation: Vapor pressure and logp. *J Chem Inf Comput Sci*. 2000;40:1046–1054.

Manuscript received May 4, 2011, and revision received Jun. 16, 2011.