Predicting Small-Molecule Solvation Free Energies: An Informal Blind Test for Computational Chemistry

Anthony Nicholls,*[,†,⊥] David L. Mobley,*[,‡,⊥] J. Peter Guthrie,[§] John D. Chodera,[#] Christopher I. Bayly,[×] Matthew D. Cooper,[×] and Vijay S. Pande[#]

OpenEye Scientific Software, Inc., Santa Fe, New Mexico 87508, Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143, Department of Chemistry, University of Western Ontario, London, Ontario, Canada N6A 5B7, Department of Chemistry, Stanford University, Stanford, California 94305, and Merck Frosst Canada Ltd., 16711 TransCanada Highway, Kirkland, Quebec, Canada H9H 3L1

Received May 10, 2007

Experimental data on the transfer of small molecules between vacuum and water are relatively sparse. This makes it difficult to assess whether computational methods are truly predictive of this important quantity or merely good at explaining what has been seen. To explore this, a prospective test was performed of two different methods for estimating solvation free energies: an implicit solvent approach based on the Poisson–Boltzmann equation and an explicit solvent approach using alchemical free energy calculations. For a set of 17 small molecules, root mean square errors from experiment were between 1.3 and 2.6 kcal/mol, with the explicit solvent free energy approach yielding somewhat greater accuracy but at greater computational expense. Insights from outliers and suggestions for future prospective challenges of this kind are presented.

Introduction

The free energy of transferring a molecule from vacuum to water, or the (aqueous) solvation free energy, plays a central role in aqueous systems. Protein folding, dynamics, and ligand binding are all closely linked to this solvation energy, because all of these processes involve exposure or protection of chemical groups from solvent. In addition, differences in solvation energy between states are required for physics-based studies of a range of other issues of interest in drug discovery including ionization states, tautomer ratios, bioactive conformations, solubility, catalysis, loop modeling, phase transfer, and aggregation and fold prediction. Thus, our ability to accurately predict absolute solvation free energies for arbitrary molecules is an essential yardstick as to how useful computational chemistry can be to drug design.

Solvation free energies also provide a good opportunity to compare methods and force fields. Other tasks (such as protein structure prediction or predicting ligand binding free energies) can require so much sampling that it can be nearly impossible to identify whether any disagreement with experiment is due to sampling, force field, or methodological problems. But because water equilibrates relatively rapidly, it is now possible to compute high *precision* estimates of solvation free energies from all-atom simulations.[1–4,58] These calculations are able to adequately sample all of the relevant degrees of freedom and so can reveal limitations in the underlying force field, when the origin of errors may remain unclear in more complicated systems. Solvation free energies also provide an opportunity to compare a wide range of different, competing approaches. Examples include solvent reaction fields applied in a self-consistent manner to Schrödinger's equation (PCM),[a,5] Poisson–Boltzmann (PB) theory,[6] group additivity methods,[7–10] and chemical typing of surface area[11] as well as various combinations thereof.[12,13]

Despite the importance of solvation free energies, accurate measurements exist for only a few hundred small molecules of limited chemical complexity and diversity. In addition, current techniques, dating back to Wolfenden and others,[14–17] require precise measurements of vapor pressure. In these experiments, molecules with large negative solvation energies have very low vacuum concentrations, making them difficult to assess. Because of this, and because of solubility issues, solvation free energies are readily available only for the range of +4 kcal/mol (insoluble) to −11 kcals/mol (strongly solvated).

The limited quantity of data, and the paucity of *new* measurements, makes solvation prediction susceptible to techniques that rely on excessive parametrization. These techniques are vulnerable to overfitting and can lack transferability and predictive power. This is a common complaint against computation in a variety of drug discovery applications: it is effective when the answers are more or less already known and ineffective when challenged with novel problems.[18,19]

Other fields have encountered similar difficulties and have ultimately benefited from testing methods via blind predictions. Research groups are asked to use their techniques on systems where experimental data are available, but withheld. Examples include the Critical Assessment of Techniques for Protein Structure Prediction (CASP),[20] Industrial Fluid Properties Simulations Challenge[21] (IFPSC), the Cambridge Crystallographic Data Centre's blind tests of small molecule crystal structure prediction,[22–24] the McMasters high-throughput screen-

* Authors to whom correspondence should be addressed [(A.N.) e-mail anthony@eyesopen.com; (D.L.M.) e-mail dmobley@gmail.com].

† OpenEye Scientific Software, Inc.

‡ University of California.

§ University of Western Ontario.

Stanford University.

× Merck Frosst Canada Ltd.

⊥ These authors contributed equally to this work.

[a] Abbreviations: HF, Hartree–Fock; AM1, Austin model 1; BCC, bond charge correction; ESP, electrostatic potential; RESP, restrained electrostatic potential fit; GAFF, generalized Amber force field; PB, Poisson–Boltzmann; PCM, polarizable continuum model; CASP, Critical Assessment of Techniques for Protein Structure Prediction; IFPSC, Industrial Fluid Properties Simulations Challenge; CAPRI, Critical Assessment of Prediction of Interactions; BAR, Bennett acceptance ratio.

10.1021/jm070549+ CCC: $40.75
Published on Web 01/24/2008

ing competition,[25] and the Critical Assessment of Prediction of Interactions (CAPRI)[26] test for protein–protein interactions. In these, the custodian releases the data only after predictions have been registered, and the successes and failures are discussed in an open forum. The value of such proceedings can be summed up by a quote from Richard Feynman: "The first principle is that you must not fool yourself and you are the easiest person to fool".[27] In a blind challenge situation there are no opportunities to adjust parameters to fit the data, rationalize, or retrospectively correct mistakes.

On the basis of these considerations, we decided to devise an informal blind challenge for solvation. Our test set was limited to 17 compounds of various complexities, and we report on the successes and failures of two methods: single-conformation PB theory and molecular dynamics free energy calculations. Both methods have little explicit parametrization for solvation, providing the opportunity to learn from failures, something not always possible with more parametrized approaches. Here, we first describe the implementations used and then present separate sets of results. Given the different viewpoints of those applying different methods here, our discussion section includes comments on each approach from each perspective.

Methods

A. The Challenge. In our challenge, one of the authors (J.P.G.) chose 17 compounds (Figure 1; Table 1) as discussed below. He obtained experimental solvation free energies for these (below), whereas the other authors made blind predictions, provided only with molecular geometries (available in the Supporting Information). Computational results presented here were calculated before experimental values were disclosed. No repeating of calculations or excluding of outliers was allowed once experimental values were revealed. The unmodified results were publicly presented in Santa Fe, NM, at the 8th Annual Customers, Users and Programmers (CUP VIII) meeting of OpenEye Scientific Software on February 27, 2007.

This informal blind test evolved, on an ad hoc basis, as a result of a conversation between J.P.G. and A.N. at the NIST workshop "Validating Modeling and Experimental Methods to Enable Drug Discovery". J.P.G. sent A.N. structures of what he considered "difficult" compounds. A.N. provided results to J.P.G., who only then forwarded experimental values. A.N. then played a role similar to that of J.P.G. for D.L.M., J.D.C., and V.S.P. A less ad hoc test would have had both groups submitting results at the same time.

B. Molecule Selection and Experimental Solvation Free Energy Determination. This test set was selected to provide a challenging test of computational methods for estimating solvation free energies. Previous testing found that additivity schemes work relatively well for solvation free energy estimates of monofunctional compounds,[7,9,10] but distant polar interactions can be very important and add additional complexities.[7,10] Thus, a challenging test set should have polyfunctional or at least highly polar molecules,[58] with a wide range of functional groups.

A test set matching this description was compiled from a variety of data sources as discussed below. A few monofunctional compounds were included for comparison with the more complicated analogues, but otherwise, as far as possible, the set is made up of molecules with interacting polar groups. The test set of compounds used in this work is given in Table 1 and Figure 1. Other forthcoming work makes use of this same set, as well as a 54 compound superset.[28]

Solvation free energies had been published for most of the 17 compounds used here, but because the titles of the papers did not contain reference to solvation, uncovering the data would have required significant determination. (Thus, we call this an "informal" blind challenge because of the distinction between availability and obscurity.)

In many cases solvation data were obtained by combining solubility and vapor pressures (publications referenced in Table 1). Vapor pressures almost invariably were obtained by extrapolation of (sometimes rather sparse) experimental data, using procedures described previously.[29,30] Phenyl formate is an exception; its solvation free energy had not previously been published. For phenyl formate, the solvation free energy is based on a new solubility determination, combined with boiling point data from the literature. Laato and Lehtonen reported boiling points at a series of pressures[31] that give a very good van't Hoff plot ($r = 0.99997$) permitting a short extrapolation to give the vapor pressure at 25 °C as 1.04 ± 0.10 Torr or 0.0014 ± 0.0001 atm. The solubility was measured by stirring the ester with 0.1 N aqueous KCl solution, centrifuging, and then adding 50 μL samples to 3.0 mL of 0.1 N aqueous NaOH and reading the absorbance due to phenoxide at 287 nm.[32] The extinction coefficient was determined by adding a suitable portion of a standard solution in acetonitrile to 0.1 N aqueous NaOH. The resulting solubility was 0.0541 ± 0.0008 M. This led to a computed free energy of solvation (relative to 1 M in gas phase and 1 M in solvent) of −4.08 ± 0.06 kcal/mol.[33]

C. Poisson–Boltzmann Calculations. The PB equation, $\vec{\nabla}\cdot[\varepsilon(\vec{r})\nabla\varphi(\vec{r})] = 4\pi\rho(\vec{r})/kT + \varepsilon(\vec{r})\kappa(\vec{r})^2 \sinh[\varphi(\mathrm{r})]$ treats water as a uniform, high dielectric (ε), responding to the electrostatic fields emanating from a charge distribution (ρ) embedded in low-dielectric solute molecules. The electrostatic potential is given by $\phi(\vec{r})$. Ionic effects can be modeled by the term $\kappa(\vec{r})^2 = 1/\lambda^2 = 8\pi e^2 I/\varepsilon\varepsilon_0 kT$, where λ is the Debye length and I is the ionic strength of the bulk solution, e is the elementary charge, and ε_0 is the dielectric of vacuum. Because water is simply represented as a uniform dielectric, with no molecular structure, this approach is often referred to as a continuum method. Effectively, solvent degrees of freedom are already integrated out, so energies include solvent entropic effects and thus are free energies. An implicit assumption is that water reorganizes more quickly than any relevant solute time scale.

Early applications of PB theory made use of analytic solutions, for example, the free energy of aqueous charging of a sphere (e.g., the Born model, Tanford−Kirkwood theory, Onsager theory, Debye–Huckel theory), cylinder, or plane (e.g., Gouy–Chapman theory). Modern use in molecular modeling dates to Warwicker and Watson[34] followed by an extensive set of studies by Honig et al.[35–39] in the mid 1980s, when computational speed enabled numerical solution of the PB equation for molecular geometries. Despite initial skepticism from the modeling community, PB has often proven to be accurate and reliable, at least for the estimation of the purely electrostatic aspects of molecular systems.[35–39]

The input to the PB equation is straightforward: the charge distribution and the dielectric as a function of atomic position. There are a variety of ways to define the dielectric function. Usually, a uniform value is assigned within the solute, and the solute boundary is determined by a hard-sphere description using standard atomic radii, such as those from Pauling or Bondi. Although standard radii are typically used, in some cases radii are treated as adjustable parameters.[42] One could avoid using atomic radii altogether by estimating the dielectric function directly from electron density, but a rigorous protocol for this is not yet known.

The other main input to the PB equation is the charge distribution. Atomic point charges are often used and can be obtained from various methods.[40,51,58] It is our experience, and that of others,[40] that charges mimicking those derived from standard RESP[64] fits to HF/6-31G* wave functions perform well with Bondi radii. One such set is the AM1-BCC formulation of Bayly et al.[59,60] Applying these charges and radii to a set of 200 small molecules with known experimental solvation energies (provided to A.N. by R. Rizzo in 2003), gives an rms error relative to experiment of 1.3 kcal/mol (data provided in Supporting Information). There is only a single adjustable or "free" parameter: a surface-area penalty designed to capture the nonpolar component of solvation. There is some debate in the literature as to how this nonpolar term should be evaluated, for instance, whether it should depend on atom type. Here, with AM1-BCC v1 charges, we use an atom-independent surface

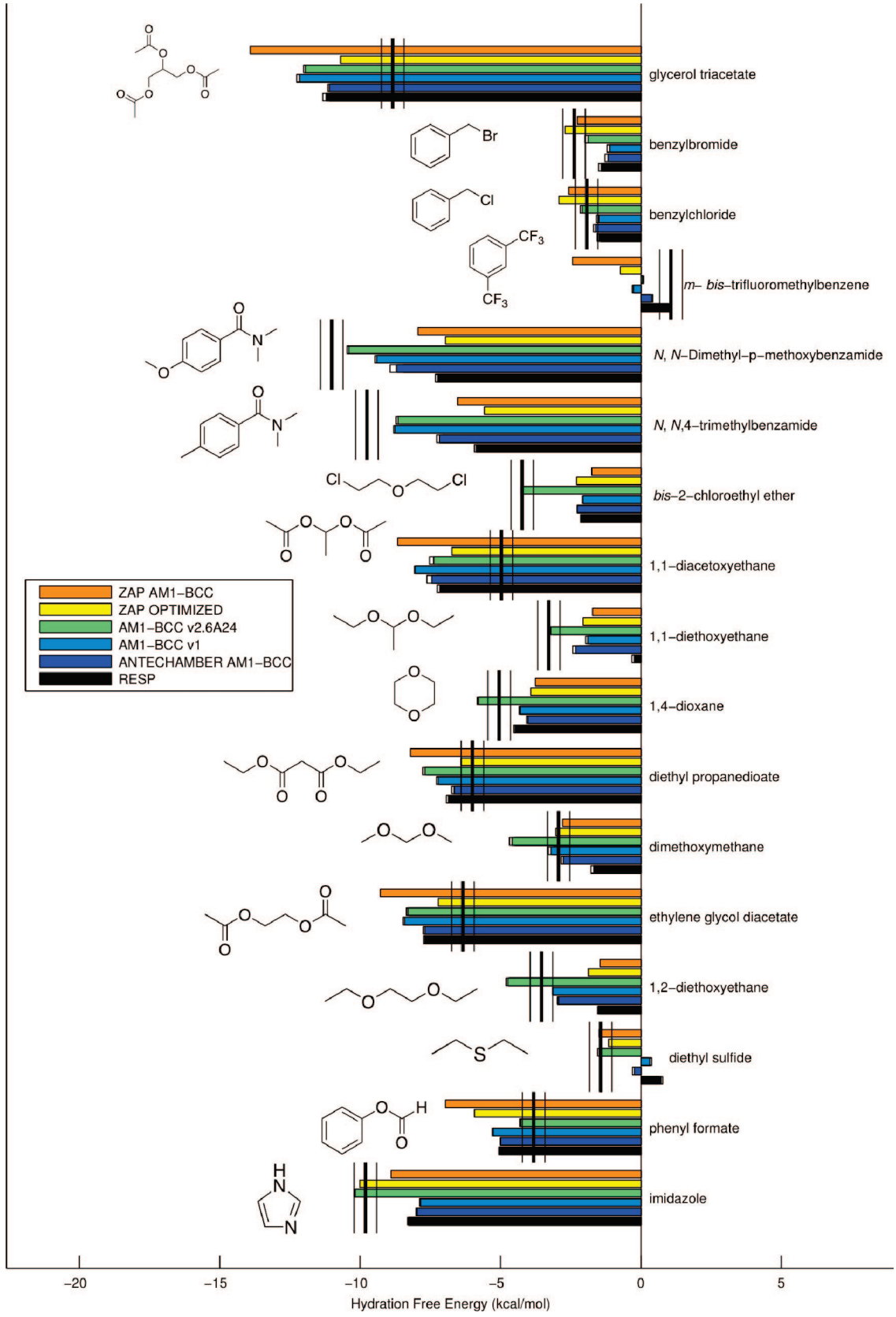

Figure 1. Solvation free energies for the 17 compounds in the challenge set computed with both alchemical free energy calculations in implicit solvent and the linearized Poisson–Boltzmann equation using ZAP. ZAP-9 denotes ZAP calculations with optimized radii. Thick vertical bars denote experimental values; thin vertical bars denote estimated experimental uncertainties.

tension of 0.01 kcal/(mol Å^2) applied to the accessible surface area (loci of centers of contact waters).

There have been a number of retrospective studies with implicit solvent models reporting rms errors of <1.3 kcal/mol, for example, with CM3, PARSE, and AGBNP. However, these models typically have 40−80 parameters.[9–11,41,42] When parametrized on only a few hundred compounds, the low data to parameter ratio raises questions of transferability. Bearing this in mind, we were interested to see

Table 1. Experimental Free Energies of Solvation[a]

no.	compound	exptl value	no.	compound	exptl value
1	glycerol triacetate	−8.84[10]	**10**	1,4-dioxane	−5.05[8,10]
2	benzyl bromide	−2.38[10,76]	**11**	diethyl propanedioate	−6.00[10]
3	benzyl chloride	−1.93[10,76]	**12**	dimethoxymethane	−2.93[8,10]
4	*m*-bis(trifluoromethyl)benzene	1.07[10]	**13**	ethylene glycol diacetate	−6.34 [10]
5	*N*,*N*-dimethyl-*p*-methoxybenzamide	−11.01[77]	**14**	1,2-diethoxyethane	−3.54[8,10]
6	*N*,*N*-4-trimethylbenzamide	−9.76[77]	**15**	diethyl sulfide	−1.43[5] (−1.55)[78]
7	bis-2-chloroethyl ether	−4.23[10,76]	**16**	phenyl formate	−3.82[79] (−4.08)[80]
8	1,1-diacetoxyethane	−4.97[10]	**17**	imidazole	−9.81[81]
9	1,1-diethoxyethane	−3.28[8,10]			

[a] Free energies (kcal/mol) for the process: gas, 1 M to aqueous, 1 M; when multiple references are given for a compound, the first is to a convenient review and the second to the source of the solvation energy. Parenthetical numbers represent potentially more accurate experimental values (J.P.G.) that were not utilized in calculating differences from experiment. The average effect of the improved numbers on rms error or any conclusions herein is small. A reasonable estimate for the experimental uncertainty is probably about 0.2 kcal/mol.[74,75]

Table 2. ZAP9 Radii

element	Bondi (Å)	ZAP-9 (Å)
carbon	1.7	1.87
hydrogen	1.2	1.1
carbonyl oxygen	1.52	1.76
secondary/tertiary nitrogen	1.55	1.40
sulfur	1.8	2.15
fluorine	1.47	2.4
chlorine	1.75	1.82
iodine	1.98	2.65

how much better PB would do with more adjustable parameters, specifically, atomic radii. Our goal was to adjust as few radii as possible and yet have an rms error similar to that of more heavily parametrized methods. The rms error from a set of experimental measurements is a function of the calculated solvation energies, which are themselves functions of the atomic radii. Using PB to calculate the gradients of the solvation energies with respect to these radii allows the rms error to be minimized by standard, gradient-driven, numerical techniques. After some experimentation as to which radii to vary, we arrived at the set listed in Table 2. With these eight radius alterations and a surface area term of 0.0064 kcal/(mol Å^2), the rms error over the Rizzo set described above was 0.80 ± 0.01 kcal/mol, comparable to the studies listed above. We termed this parametrization ZAP-9. When the single heavy atom compounds (water, ammonia, hydrogen disulfide, and methane) were removed from the data set, the rms error dropped to 0.76 ± 0.01 kcal/mol. On further examination, we observed that acetamide, *N*-methylacetamide, 2-methylpyrazine, and 2-ethylpyrazine were the most significant outliers, accounting for almost 20% of the total variance. This data set is provided in the Supporting Information.

The PB approach to the blind-challenge set was straightforward. We used starting conformations generated by J.P.G. These had been initially prepared using PCModel.[43] When multiple conformations were expected, the lowest energy conformation generated by GMMX in PCModel was used. These structures were then geometry optimized using Gaussian03[44] at the B3LYP/6-31G** level of theory in vacuum. Examination suggested these structures were reasonable; however, choosing to use a single conformation to predict solvation raises two important issues. The first is whether the correct, that is, lowest energy, conformation is used. As the lowest energy conformation in vacuum may not be the lowest in solvent, the transfer step may formally require a change in conformation. However, it has been our observation that solvation flattens the conformation energy landscape,[45] causing this to be less of a problem than might be assumed. However, this very flattening also leads to a conformational entropy component to the solvation free entropy of transfer. This issue and the potential for conformation dependence and conformational change will be discussed elsewhere.[46]

With these starting structures, we computed charges with the OpenEye implementation of AM1-BCC v1,[59,60] without additional geometry optimization. We then ran ZAP v2.0, OpenEye's PB solver,[47] using the ZAP-9 radii, with a relative dielectric of 80 for water and unity for the internal dielectric. (Although this is not a physically correct dielectric for organic molecules, we have observed that AM1-BCC charges work well with this assignment.[45] This could, of course, be considered an implicit parameter.) We used a grid resolution of 0.5 Å and a buffer region around the molecule of 2 Å from surface to box boundary. Due to grid discretization, the estimated precision of such calculations is ∼0.1 kcal/mol. Computing charges for the 17 molecules took a total of 0.31 s, and running ZAP took another 0.25 s on a 3 GHz Intel processor.

D. Alchemical Free Energy Calculations. Alchemical free energy calculations provide a rigorous way to calculate free energy differences between thermodynamic states (for example, a state with a solute in water versus a state with water plus a separate solute in vacuum), within classical statistical mechanics, assuming the underlying force field accurately describes the physics. For recent reviews, see refs 48 and 49. These calculations can in principle be performed with either explicit solvent or implicit solvent, such as the PB methods mentioned above. Whereas it is most common to use alchemical methods with explicit solvent, alchemical methods coupled to implicit solvent would provide a correct way to include the missing conformational entropy of the solute and other factors, while avoiding the computational expense of a full explicit solvent calculation.[46]

In these alchemical calculations, free energy differences are computed by mutating the system from an initial thermodynamic state (here, the solute in a bath of water equilibrated at 1 atm and 300 K) to a final state (the empty water bath with the solute removed plus the solute in vacuum) by a series of alchemical intermediates. These intermediates are introduced to ensure the configurations sampled in neighboring states are not too different (or, more precisely, that there is adequate phase-space overlap), a requirement for precise free energy differences.[50,51] A second requirement is that all relevant system conformations must be sampled in each state.[52,53] The first requirement can be tested using error analysis and experimentation (i.e., by testing different numbers of intermediates). The second, convergence, can be assessed by examining estimates from multiple initial geometries, as was done here. In the limit of sufficient sampling, results obtained from alternate starting conformations will agree to within statistical uncertainty.[52]

In this work, molecular dynamics simulations were done with the GROMACS 3.3[54,55] simulation package, using critical bug fixes described previously.[58] TIP3P solvent[56] was used for the solvent bath, and small-molecule parameters were taken from the Amber GAFF[57] force field.

Several charge models were tested, with the expectation that free energies computed using charges from multiconformer RESP fits would be the most accurate, as suggested by a previous retrospective study using the same protocols but with single-conformer RESP fits.[58] The charge models considered were AM1-BCC version 1,[59,60] both as provided by Christopher Bayly[61] and as computed by ANTECHAMBER[62] version 1.2.4 using the top single conformer generated by the conformer enumeration program Omega, version 2.1.0, from OpenEye (with default settings). We also obtained partial charges computed with AM1-BCC v2.6A24, currently under

Table 3. rms Errors Relative to Experiment for Explicit Solvent Charge Models

charge model	rms error (kcal/mol)
Alchemical Free Energy Calculations	
multiconformer RESP HF/6–31G*	2.05 ±0.05
Merck-Frosst AM1-BCC v1	1.71 ± 0.05
ANTECHAMBER AM1-BCC v1	1.53 ± 0.05
Merck-Frosst AM1-BCC v2.6A24	1.33 ± 0.05
Single-Conformer Poisson–Boltzmann	
PB: OpenEye-AM1-BCC v1, ZAP-9	1.87 ± 0.03
PB: OpenEye-AM1-BCC v1, Bondi radii	2.57 ± 0.03

development, by Christopher Bayly.[63] Finally, we computed multiconformer RESP fit charges.[64] For each molecule, a number of conformers were automatically generated with Omega (as described above). The geometry for each conformer was optimized with HF/6-31G* in GAMESS version 7 Sep. 2006 (R4) using the quadratic approximation augmented Hessian minimizer. Optimization was terminated when the gradient was $<10^{-4}$ hartree/bohr or 100 steps were taken. We computed the electrostatic potential on a series of surfaces at 1.4, 1.6, 1.8, and 2.0 times the van der Waals radii from the atomic nuclei with a grid spacing of 1 $Å^{-2}$. Charges were fit to the electrostatic potential for all of the conformations simultaneously with two-stage process using ANTECHAMBER 1.27 and the associated "respgen" program (version 2.1). All conformers were weighted equally in the fit.

Simulation protocols were as described previously,[58] except as follows. First, instead of conducting only one set of simulations at each alchemical intermediate (λ value), we checked convergence by running three separate trials from different starting conformations generated using Omega. For some molecules, Omega predicted only a single conformer, in which case we ran three trials initiated from the same conformer. Additionally, in this study, following an initial equilibration period (110 ps at each λ value, as described previously), production simulations (from which data were collected) were run for 10 ns at each λ value rather than 5 ns. Finally, we performed separate solute electrostatic annihilation calculations for each charge model, rather than computing the free energy of changing the charges from a reference set as in the previous work.

Free energies were computed using the Bennett acceptance ratio method (BAR),[65,66] as described previously.[58] BAR has been shown to provide a maximum-likelihood estimate of free energy differences in alchemical free energy calculations and appears to be substantially more efficient in a number of tests[50,67] than alternatives such as exponential averaging or thermodynamic integration.

Error analysis was done by autocorrelation analysis and block bootstrap as in previous work.[58] We also compared the standard deviation across the three trials for each molecule with these computed uncertainties. Although the use of three trials provides only a crude estimate of the standard deviation, it was always at least as large as the uncertainties computed from the block bootstrap procedure, and sometimes several times larger, suggesting the block bootstrap analysis might underestimate the true uncertainty. Therefore, we opted to report the three-trial standard deviation as the more conservative uncertainty estimate.

Timings varied substantially due to variations in system size. For each molecule, a solvation free energy calculation took 1–18 h with 21 AMD Opteron processors (2.5 GHz). Additional charge sets required only a discharging calculation, requiring only five processors for a similar length of time.

Results

A. Poisson–Boltzmann. The rms error for PB (ZAP-9, AM1-BCC v1) over the test set was 1.87 ± 0.03 kcal/mol (Table 3; Figure 2). The benzamides were significant outliers and accounted for over half of the variance. Without these, the rms error would have been 1.3 ± 0.01 kcal/mol. These errors are larger than those from the more extensive Rizzo training set (rms error = 0.76 ± 0.01 kcal/mol). However, the challenge set was intentionally more difficult. Also, acetamides had relatively large errors in the Rizzo set, forecasting significant errors for challenge-set benzamides. In fact, consistency of outliers was a positive feature, encouraging investigation of underlying causes.

In ZAP-9, certain types of radii (Table 2) are adjusted to improve agreement with experimental hydration free energies. Treating radii as adjustable parameters can certainly improve the overall agreement with experiment but may actually introduce problems for specific molecules. Acetamide and *N*-methylacetamide from the Rizzo set and the benzamides from the challenge set provide an interesting case study of this phenomenon. In ZAP-9, the optimization of radii against the Rizzo set expands the radius of carbonyl oxygens and reduces the nitrogen radius for secondary and tertiary amines. To understand this, we compared AM1-BCC v1 charges with the HF/6-31G* RESP-fit charges for carbonyls. Charges derived from HF/6-31G* have long been used for aqueous-phase molecular simulations because such wave functions typically give overpolarized charge sets, supposedly mimicking the polarizing effects of water.[58] AM1-BCC charges are generally similar because they are parametrized to HF/6-31G*-derived charges, supplemented by alchemical free energy calculations.[59,60] However, we found that the carbonyls are, in general, overpolarized even compared to HF/6-31G* RESP charges. This is particularly true for aldehydes, causing substantial disagreement between experimental gas-phase dipole moments and those predicted by AM1-BCC and generating PB-derived solvation energies too negative compared to experiment. The ZAP-9 radii optimization counters this by enlarging the carbonyl oxygen, effectively pushing water further away from the charge on the oxygen. However, acetamides are an exception. Here, AM1-BCC v1 produces charges very close to those from HF/6-31G* RESP. Therefore, the global radius adjustment for carbonyl oxygens leads to an unnecessary change to the benzamide carbonyl group. We estimate this leads to an underestimate of the energetic preference for water of about 1.5 kcal/mol, given the average behavior of aldehydes in the Rizzo set relative to experiment (although the error for benzaldehyde was closer to 2.5 kcal/mol).

The second aspect of error for the benzamides is the tertiary nitrogen. In ZAP-9, secondary and tertiary nitrogens are made smaller to mimic the "methylamine" anomaly: Adding a methyl group to ammonia actually increases solubility, and the change in solvation from methylamine to di- and trimethylamine is surprisingly small.[72] Bayly et al. explicitly recognized this fact by increasing the bond charge increment parameters in AM1-BCC v1 for such nitrogens.[59,60] Although this may improve explicit solvent results, it is insufficient to enable PB to reproduce the experimental trend. The reduction of the secondary and tertiary nitrogen radius in ZAP-9 increases the solvation energy appropriately, that is, allowing water closer to the nitrogen charge. However, as Figure 3 shows, this change affects different classes of molecules in different ways, depending crucially on the local environment. The solvation contribution from a relatively exposed nitrogen is sensitive to radius, but that from the inaccessible nitrogens in the challenge-set benzamides is not. Essentially, the tertiary nitrogen radius is the wrong variable to adjust, and the benzamides remain uncorrected. Unfortunately, there is no simple insight into what the right variable might be for the methylamines, but by looking at other secondary and tertiary nitrogens in the Rizzo set it is possible to estimate a necessary correction factor of between 1.5 and 2.0 kcal/mol. Taken together, parametrization issues for the carbonyl oxygens and tertiary nitrogens yield an expected error for benzamides of between 3.0 and 4.5 kcal/mol, consistent with the ~4 kcal/mol observed.

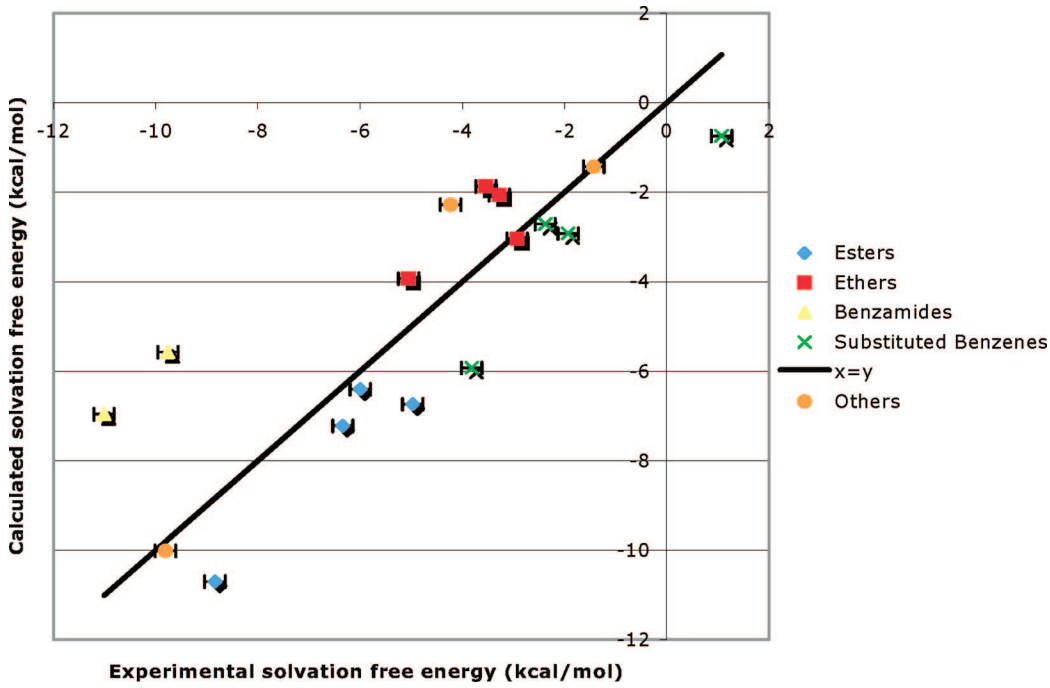

Figure 2. Solvation energies with ZAP, AM1-BCC v1.0 without geometry optimization, and radii and nonpolar contribution as in the ZAP-9 parametrization. Expected precision in the calculations is about 0.1 kcal/mol. Molecules are color coded by group as described in the text and legend. Experimental error bars are 0.2 kcal/mol.

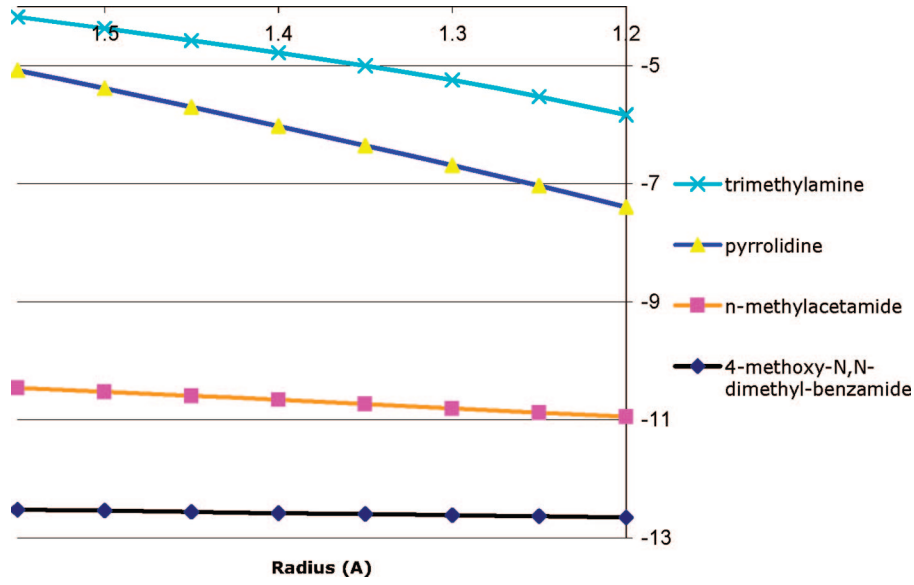

Figure 3. Solvation energies, from ZAP, as a function of the amine nitrogen radius. This illustrates the problems of parametrizing with the wrong variable. Whereas the observed effect on solvation is relatively independent of chemical environment, the PB energy is very sensitive. In particular, the solvation of the benzamides in the challenge set is almost completely insensitive to changes in the radius of the tertiary nitrogen.

B. Explicit Solvent. We estimated solvation free energies in explicit solvent using several charge models, with the Amber GAFF forcefield for bonded and Lennard-Jones parameters. Previous work tested the accuracy of different charge models for solvation free energies with the GAFF parameters and found that HF/6-31G* RESP charges and AM1-BCC v1 charges give fairly good accuracies (slightly below 1 kcal/mol rms error), whereas charges fitted to potentials calculated using higher levels of quantum theory did not appear to yield a marked improvement, and accuracies were only comparable when a reaction field was employed.[58] This might suggest that beyond a given level of quantum theory polarization of the solute becomes more important than an accurate charge distribution or that the water model or Lennard-Jones parameters, rather than electrostatics, are limiting accuracy. Alternatively, this could be the accuracy limit for atom-centered point-charge electrostatics.

In any case, we focused our efforts here on RESP charges and several different variants of AM1-BCC. As the solute molecules considered in the previous study were mostly rigid with few rotatable bonds, whereas the molecules in the current study typically had multiple rotatable bonds, we performed multiconformer RESP fits (as described under Methods) to ensure the resulting charge set was not overly biased by a single conformation. Resulting rms errors for the charge models considered here are shown in Table 3.

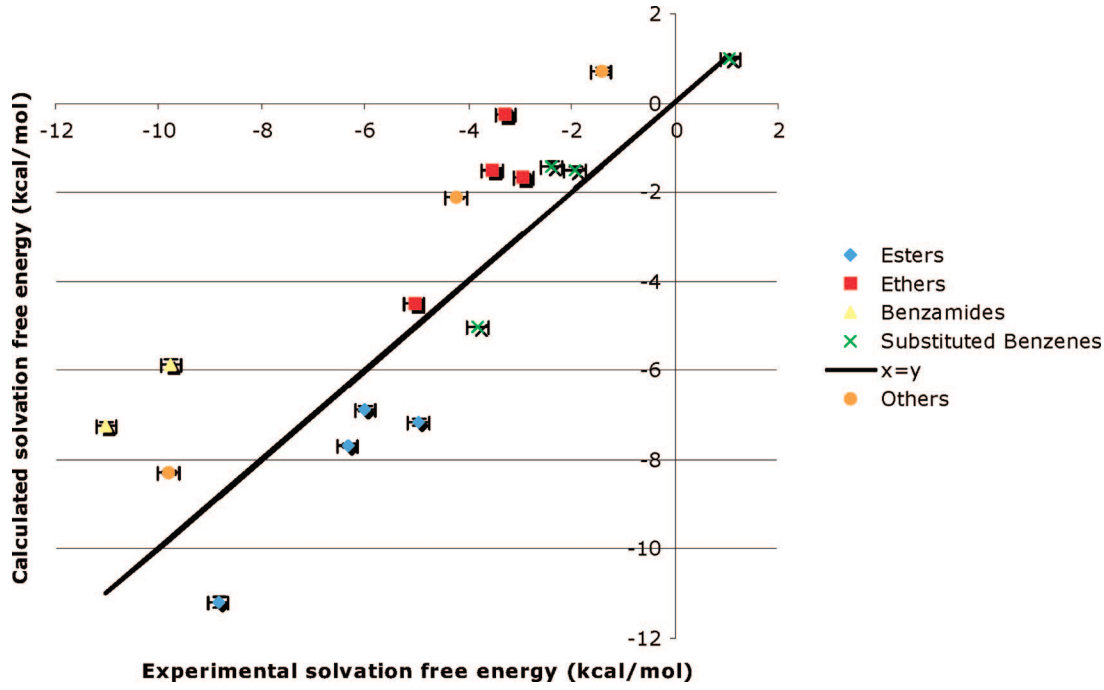

Figure 4. Solvation free energies in explicit solvent, calculated with alchemical free energy calculations and RESP charges, versus experiment. Molecules are color coded by group as described in the text and legend. Vertical (computed) error bars are one standard deviation; horizontal (estimated experimental) error bars are 0.2 kcal/mol.

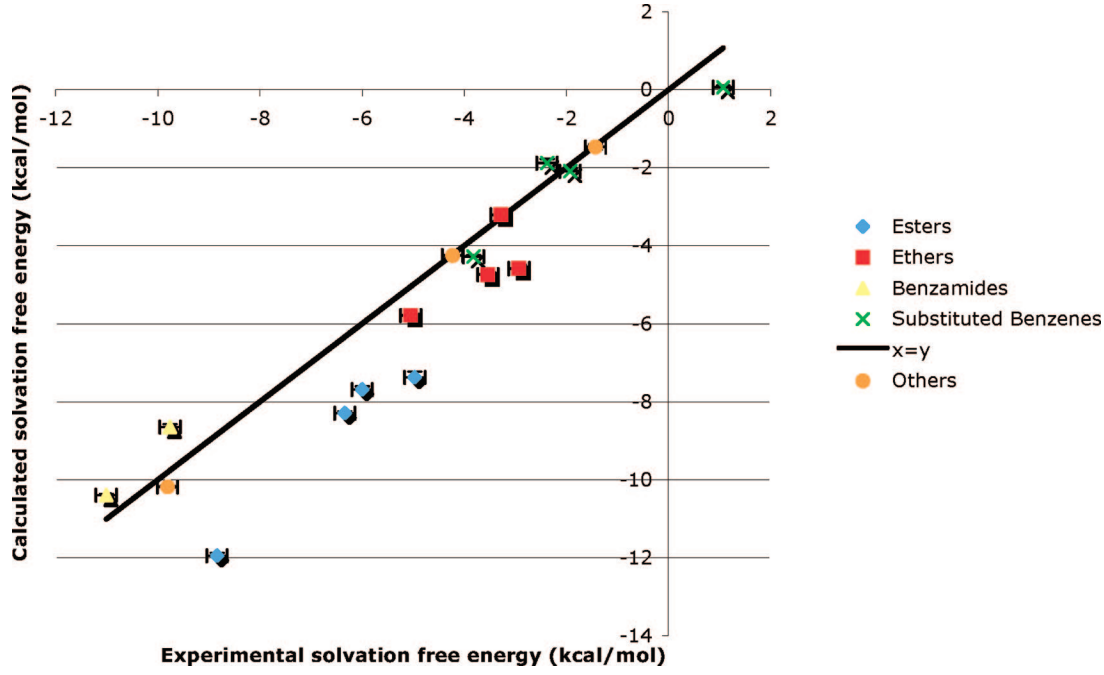

Figure 5. Solvation free energies in explicit solvent, calculated with alchemical free energy calculations and AM1-BCC v2.6A24 charges, versus experiment. Molecules are color coded by group as described in the text. Vertical (computed) error bars are one standard deviation; horizontal (estimated experimental) error bars are 0.2 kcal/mol.

In these prospective studies, the new version of AM1-BCC currently under development (v2.6A24) performed best (rms error = 1.33 ± 0.05 kcal/mol), followed by AM1-BCC v1 (rms error = 1.71 ± 0.05 to 1.53 ± 0.05, depending on the implementation). To our surprise RESP HF/6-31G* charges actually performed worst (rms error = 2.0 ± 0.2 kcal/mol).

Figures 4 and 5 show results with AM1-BCC v2.6A24 and RESP charges, color-coded by chemical group. Compounds are grouped into esters (of which there were four), ethers (four), substituted benzenes (four), benzamides (two), and others (three). It is interesting to note that, for both of these charge models, the esters as a group appear to have solvation free energies that are excessively hydrophilic. For AM1-BCC v2.6A24, if the esters are excluded, the rms error drops from 1.33 ± 0.05 to 0.78 ± 0.06 kcal/mol. If, instead of the esters, the next worst four compounds are excluded, the rms error is 1.40 ± 0.06 kcal/mol, indicating that the esters really are the worst outliers as some of these next four are more accurate than the mean.

Our previous retrospective study of explicit solvent solvation free energies[58] included no esters. However, we can compare ANTECHAMBER AM1-BCC v1 results here with those from the previous study for other classes of compounds. The previous set included one ether, with an error of 0.7 kcal/mol (computed free energies were too positive), consistent with the rms error on ethers here (0.8 kcal/mol, too positive). The set also included 4-methylimidazole, similar to imidazole; the two have comparable errors (1.9 versus 1.8 kcal/mol; too positive). The error on ethyl methyl sulfide was 1.2 kcal/mol and that on diethyl sulfide, 1.2 kcal/mol (both too positive). The retrospective set also included benzamide, with an error of 1.1 kcal/mol, higher than average. Here, the benzamides also have higher errors than average. Again, computed values are too positive. Thus, it seems that errors for particular chemical classes in our previous work track well with errors for particular classes in this study. Thus, because esters are particularly problematic here, it seems likely that systematic errors for esters will become clear in future studies.

Discussion

The results presented above are, we believe, the first reported *prospective* test of our respective methods for calculating solvation free energies. Both methods performed well, although the errors were larger than either group expected on the basis of previous work. This is typical of prospective studies, indicating that our methods are perhaps not immune from the problems of extrapolation seen in highly parametrized methods,[11–13] although the test set here was chosen to be more challenging than typical training sets.

This test provided a great deal of insight into remaining deficiencies. In particular, compounds most poorly predicted provided the most insight. Overall, explicit solvent alchemical free energy calculations gave the lowest errors relative to experiment, but at considerably more computational expense. One would hope that including full atomic detail would make the explicit solvent calculations more accurate, and this was indeed the case. That PB does as well as it did is, to some extent, surprising. This test provided an opportunity to honestly consider each method's current strengths and weaknesses. Below we summarize key lessons from this study, what could be done better, and the challenges inherent in making prospective studies more common.

PB. Encouragingly, it was relatively easy to understand PB failures. The benzamides led us to a much better appreciation of the nature of the AM1-BCC v1 charge set and of the ultimate futility of attempting to parametrize one set of physical variables (radii) to compensate for variance in another (charges) (see Figure 3). In one sense, radii parametrization was successful, improving the prospective results (ZAP-9 rms error of 1.87 ± 0.03 kcal/mol vs 2.57 ± 0.03 kcal/mol using Bondi radii), but it also contributes to the magnitude of the benzamide failure by incorrectly increasing the radius of the carbonyl oxygen.

Closer investigation of the AM1-BCC v1 bond charge increment parameters revealed that aldehydes, the root cause of the carbonyl radii parametrization issues in ZAP-9, are actually treated differently from other carbonyl-containing groups. Were this difference in parametrization removed, PB results would improve, potentially reducing the need for compensatory parametrization of radii. However, Bayly et al. parametrized AM1-BCC v1 to both HF/6-31G* RESP charges and to reproduce relative solvation energies of isosteric groups in explicit solvent, so it seems reasonable to assume this parametrization of oxygen types was driven by the explicit solvent results. Indeed, AM1-BCC v1 performs quite well in the explicit solvent simulations presented here, supporting this conclusion. Thus, charge schemes derived for use with explicit solvent simulations may only incidentally be useful for PB calculations.

That AM1-BCC v1 charges do perform well with PB is probably because they are *predominantly* parametrized to vacuum electrostatic potentials rather than explicit solvent simulations. We found it significant that the rms errors for AM1-BCC v1 were very similar between free energy calculations in explicit solvent and PB with radii optimization (1.71 ± 0.05 vs 1.87 ± 0.03 kcal/mol). Apparently, radii can compensate for charges that have been tuned for explicit solvent, but radius adjustment has only a limited ability to capture such effects.

Two clear directions for improvement have emerged from this work. First is to build a charge set more appropriate for PB. It would be elegant if this were also consistent with experimental electrostatic observables such as gas-phase dipoles and quadrupoles. These are not widely used as a reference set for modeling because water polarizes charge distributions away from those appropriate for vacuum. However, a solute dielectric function in theory captures this polarization.[48,68] A potential approach would be to use dielectric values derived from experiment in conjunction with point charges, or possibly higher order multipole moments, from quantum mechanics at a sufficient level to reproduce gas-phase moments. The other direction is a more satisfactory continuum theory of the nonpolar term, something explicit solvent simulations can fairly claim to capture more accurately.

Explicit Solvent. The explicit solvent free energy calculations discussed here take a very different approach from the PB calculations. One appealing feature of the free energy calculations is their ability to rigorously estimate the correct solvation free energies for an underlying force field, within classical statistical mechanics. Of course, because that force field is not perfectly representative of reality, these may still disagree with experiment. In contrast, PB calculations neglect some of the physics present in the explicit solvent simulations (such as solvent granularity, long-range dispersion interactions, solute entropic factors, and other statistical mechanical issues[46]). Thus, explicit solvent calculations could be expected to give somewhat higher accuracy. Indeed, they did perform slightly better than PB, with rms errors ranging from 1.33 ± 0.05 to 1.71 ± 0.05 kcal/mol with different AM1-BCC charges and slightly higher for RESP charges. The best PB results, with adjustable radii, came in at 1.87 ± 0.03 kcal/mol for this set. Therefore, on the one hand, it is encouraging that explicit solvent free energy calculations did perform better than PB. On the other hand, the fact that the difference was not more pronounced provides a challenge for future explicit solvent simulations, and suggests several avenues for further exploration. Can implicit solvent be made as accurate as explicit solvent while remaining less computationally expensive? Or can explicit solvent be made substantially more accurate with physics-based improvements? Future studies will be necessary to answer these questions.

Performance in this prospective test was substantially worse than that reported in a previous study using very similar methods.[58] Why was the performance so much worse? The force fields used here are not parametrized for solvation free energies, except the very few solvation-related parameters used in the AM1-BCC charge models, so there is no reason in principle to expect prospective studies to be any more difficult than retrospective studies. However, previous retrospective work[58]

suggested that with increasing polarity, errors in computed solvation free energies tend to grow rapidly. This test set contains a larger fraction of quite polar compounds, so an rms error only twice as large as that in the previous study may actually be surprising.

These explicit solvent free energy calculations highlight a number of potential parameter issues. In particular, solvation free energies for esters were quite inaccurate with every charge model tested. Benzamides and imidazole derivatives also appeared to be problematic in this and the previous retrospective work, although not as problematic as with implicit solvent.

Another issue is polarization. Fixed-charge force fields may not yield accurate solvation free energies, because they neglect the energy cost of polarizing a molecule's charge distribution from the vacuum distribution to the solvent distribution. Previous estimates of this cost indicate it can be significant for some compounds.[58,69]

Explicit Solvent Simulations from a PB Perspective. An interesting symmetry exists between the approaches in that neither currently includes polarization, and both use an internal dielectric constant of unity. Yet both appear to be capable of capturing a good fraction of the physics of solvation. The charge sets used, with their overpolarization relative to gas-phase charges, clearly perform fairly well, although there may eventually be instances when this overpolarization causes problems. It is worth also noting that AM1-BCC v2.6A24 has eight additional parameters designed to improve explicit water simulation estimates of solvation for certain classes of compounds. Similar "solvation" adjustments are to be found in the CM3 charge model from Cramer et al.[70] and the van der Waals parameters of OPLS by Rizzo and Jorgenson.[71] It will be interesting to compare both explicit solvent free energy methods and PB using AM1-BCC v2.6A24. Comparison at the AM1-BCC v1 level shows both approaches have similar rms errors, but this is perhaps unfair to explicit solvent simulations as ZAP-9 has nine "solvation" parameters. Even though some of these parameters are needed just to adapt AM1-BCC to continuum electrostatics, a fair appraisal would give the edge to explicit solvent simulations.

Finally, the PB solvation estimates are still much faster than explicit solvent simulations. Does this matter? The timings for explicit solvent simulations are not unreasonable, and were the error consistently sub-kilocalorie per mole, the approach might well be worth the effort for applications where solvation is a dominant component, for instance, in predicting tautomer ratios. However, both methods still have considerable room for improvement, making future tests all the more interesting. Such work may also produce mutual benefit, with explicit solvent simulations providing calibration for some aspects of continuum theory such as nonpolar components, and continuum methods providing useful insight into polarization, ionic effects, and the behavior of macroscopic systems.

PB from an Explicit Solvent Simulation Perspective. PB is a simple, computationally efficient solvent model. As applied in this work, it neglects certain aspects of the underlying physics of solvation that are undoubtedly important in some situations. One example is solute entropic effects. In the gas phase a molecular energy landscape is typically very rough, because electrostatic interactions are unshielded, and a molecule may have a single dominant conformation. In contrast, solvent provides electrostatic shielding, smoothing energy landscapes and allowing a wider potential range of conformations. For molecules with a significant number of rotatable bonds (of which there are examples in this test set) this may result in significant entropy gains on transfer to water. Nothing about PB itself prohibits these effects from being included. Instead of estimating solvation energies from single conformations, or averaging over conformations, one would derive solvation energy estimates using statistical mechanics. This presents a promising avenue for future work[46] and could bridge part of the gap between the PB and the explicit solvent free energy approaches presented here.

Other differences between PB and explicit solvent also merit further investigation. The PB treatment of the nonpolar component of solvation here is extremely simple. Alchemical free energy calculations provide a way to calculate this nonpolar component from explicit solvent simulations; some data on this are already available.[2,58] These can be used to improve the treatment of nonpolar solvation in implicit solvent models beyond the simple surface area approximation. Solvent granularity effects may also be important on small length scales as in, for example, the asymmetry of water with respect to positive and negative charges,[72] surface curvature,[73] and solvent-separated ion pairs. Again, explicit solvent free energy calculations can be used to help understand this and potentially improve implicit solvent models.

It is intriguing that, in view of the physics missing from PB calculations, rms errors in solvation free energies from PB and all-atom simulations are relatively similar. This could suggest that both approaches are limited by some common factor, possibly the neglect of polarization, whereas better treatment of dispersion and solute entropies may be why explicit solvent simulations perform better.

What We Could Do Better. Because this exercise evolved rather than having been designed from the outset, the selection of test compounds could have been more representative of drug-like molecules, although it turned out that the test set of 17 compounds used spanned the full range of solvation energies in the larger J.P.G. set of 54 compounds. The test set was overweighted in esters, and too many of these had similar solvation energies. However, the functional groups represented in compounds for this test (amides, polyfluoro compounds, haloethers, polyesters) will continue to be good candidates for future tests. Other functional groups found in drug-like molecules should be added as data become available, including sulfonyl derivatives, phosphoryl derivatives, nitro compounds, and heterocycles, to name just a few. Clearly a larger test set would help, although diversity is also important.

Only by testing methods with polyfunctional molecules with strongly interacting functional groups will it be possible to test whether any parametrization is adequate to cover the full range of drug-like molecules, and if not, to drive improvements in the underlying methods and force fields. Such developments are unlikely to happen until theory is confronted with a challenging set of compounds. Unfortunately, it is difficult to obtain such data.

In future blind tests, we hope to have other groups apply different methods. Blind testing a greater diversity of techniques can only provide more insight. Thus, publicizing the next challenge, and issuing specific invitations, seems to be worthwhile.

Conclusions

The results presented above are for a small set of compounds and may not be representative of the results for a larger set or for larger molecules. The average number of non-hydrogen atoms here was about nine, less than half that of a typical drug. However, the accuracy of both methods appears to be relatively insensitive to molecular size, and to an extent solvation is an

additive phenomenon, at least in the absence of distal polar interactions.[7,10] Therefore, there is no intrinsic reason drug-like molecules could not be handled effectively. But are calculations reliable for novel molecules? The results here are at worst promising and at best exciting. If, with improvements, these methods can begin to reliably predict solvation, then more complex phenomena may become computationally tractable. This may have immediate impact on important yet experimentally challenging properties such as tautomer ratios. It is tempting to view the free energy calculation results as evidence we have arrived; that is, fix the esters, and the goal is achieved. However, medicinal chemistry is so vast and varied, one cannot be sure the next set of compounds will not reveal the need for further adjustment, followed by then another for the next set and so on. To truly have confidence in computational methods, regular repetition of this process is essential. For this we need more solvation data, whether from the literature or new experimental measurements, although such are rare.

The apparent lack of experimental attention to this important property appears, in retrospect, surprising, but it can also be seen as an opportunity. Other fields have progressed rapidly whenever experimentalists and theoreticians have actively collaborated. Solvation energies provide a similar opportunity, one that computational chemistry should not ignore. Although there are hundreds of solvation energies in the literature, most of them are for monofunctional compounds rather than the polyfunctional compounds that provide the more challenging test for theory. In addition, there are numerous polar functional groups either not represented in the published data or represented by only one or two compounds. In either case, measuring the needed solvation energies will pose experimental challenges. For highly soluble compounds the problem is to detect the tiny amount of material in the gas phase at equilibrium, whereas for sparingly soluble compounds with quite negative free energies of solvation, the problem is to measure the very low vapor pressure. There is a need for creative solutions to these analytical problems.

Acknowledgment. We thank Christine M. Isborn (University of Washington) for help with GAMESS calculations and Christopher I. Bayly (Merck-Frosst) for providing several versions of AM1-BCC charges. D.L.M. thanks Ken A. Dill for support. A.N. thanks Robert C. Rizzo (State University of New York at Stony Brook) for his curation of the solvation dataset used in the ZAP-9 parametrization.

Note Added after ASAP Publication. This manuscript was released ASAP on January 24, 2008, with an incomplete listing of authors and affiliations. The correct version was posted on February 6, 2008.

Supporting Information Available: The 17 structures forming the prospective data set; structures for the training set for ZAP-9, with charges and radii; raw data for the alchemical free energy calculations, including the total and average charging and nonpolar components for each compound; computed hydration free energies with each method. This information is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. 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*, 5740–5761.

(2) Shirts, M. R.; Pande, V. S. Solvation free energies of common amino acid side chain analogs for common molecular mechanics water models. *J. Chem. Phys.* **2005**, *122*, 134508–134513.

(3) Hess, B.; van der Vegt, N. F. A. Hydration thermodynamic properties of amino acid analogues: A systematic comparison of biomolecular force fields and water models. *J. Phys. Chem. B* **2006**, *110*, 17616–17626.

(4) Deng, Y.; Roux, B. Hydration of amino acid side chains: nonpolar and electrostatic contributions calculated from staged molecular dynamics free energy simulations with explicit water molecules. *J. Phys. Chem. B* **2004**, *108*, 16567–16576.

(5) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3094.

(6) Jean-Charles, A.; Nicholls, A.; Sharp, K.; Honig, B.; Tempczyk, A.; Hendrickson, T. F.; Still, W. C. Electrostatic contributions to solvation energies: comparison of free energy perturbation and continuum calculations. *J. Am. Chem. Soc.* **1991**, *113*, 1454–1455.

(7) Hine, J.; Mookerjee, P. K. the intrinsic hydrophlic character of organic compounds. Correlations in terms of structural contributions. *J. Org. Chem.* **1975**, *40*, 292–298.

(8) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *J. Soln. Chem.* **1981**, *10*, 563–595.

(9) Guthrie, J. P. A group equivalents scheme for free energies of formation of organic compounds in aqueous solution. *Can. J. Chem.* **1992**, *70*, 1042–1054.

(10) Guthrie, J. P. Concerning the distant polar interaction in free energies of transfer: an explanation and an estimation procedure. *Can. J. Chem.* **1991**, *69*, 1893–1903.

(11) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. SM6: a density functional theory continuum solvation model for calculating aqueous solvation free energies of neutrals, ions, and solute-water clusters. *J. Chem. Theor. Comput.* **2005**, *1*, 1133–1152.

(12) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. Class IV charge models: a new semiempirical approach in quantum chemistry. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 87–110.

(13) Gallicchio, E.; Yu, L.; Levy, R. M. The SGB/NP hydration free energy model based on the surface generalized born solvent reaction field and novel nonpolar hydration free energy estimators. *J. Comput. Chem.* **2002**, *23*, 517–529.

(14) Wolfenden, R.; Andersoon, L.; Cullis, P. M.; Southgate, C. C. B. Affinities of amino acid side chains for solvent water. *Biochemistry* **1981**, *20*, 849–855.

(15) Cabani, S.; Conti, G.; Lepori, L. Thermodynamic study on aqueous dilute solutions of organic compounds. Part 1. Cyclic amines. *Trans. Faraday Soc.* **1971**, *66*, 1933–1942.

(16) Abraham, M. H.; Andonianhaftvan, J.; Whiting, G. S.; Leo, A.; Taft, R. S. Hydrogen bonding. Part 34. The factors that influence the solubility of gases and vapours in water at 298 K, and a new method for its determination. *J. Chem. Soc., Perkin Trans. 2* **1994**, *2*, 1777–1791.

(17) Betterton, E. A.; Hoffmann, M. R. Henry's law constants of some environmentally important aldehydes. *Environ. Sci. Technol.* **1988**, *22*, 1415–1418.

(18) Kubinyi, H.; Hamprecht, F. A.; Mietzner, T. Three-dimensional quantitative similarity−activity relationships (3D QSAR) from SEAL similarity matrices. *J. Med. Chem.* **1998**, *41*, 2553–2564.

(19) Golbraikh, A.; Tropsha, A. Beware of q2. *J. Mol. Graph. Mod.* **2002**, *20*, 269–276.

(20) Protein Structure Prediction: A Special Issue. *Proteins: Struct., Funct. Genet.*, **1995**, 295–460.

(21) IFSPC: www.fluidproperties.org.

(22) Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. A test of crystal structure prediction for small organic molecules. *Acta Crystallogr.* **2000**, *B56*, 697–714.

(23) Motherwell, W. D. S. Crystal structure prediction of small organic molecules: a second blind test. *Acta Crystallogr.* **2002**, *B58*, 647–661.

(24) Day, G. M.; et al. A third blind test of crystal structure prediction. *Acta Crystallogr.* **2002**, *B61*, 511–527.

(25) McMaster University Data-Mining and Docking Competition. *J. Biomol. Screen.* **2005**, *10*, 647–686. (special issue).

(26) CAPRI: Critical Assessment of Predictions of Interactions: www.capri.ebi.ac.uk.

(27) Commencement address to the 1974 Caltech graduating class.

(28) Guthrie, J. P.; Povar, I. Manuscript in preparation.

(29) Guthrie, J. P.; Taylor, K. F. Additivity methods for estimating the heat-capacities of vaporization for organic compounds. *Can. J. Chem.*, **1984**, *62*, 363–372,

(30) Guthrie, J. P. Estimation of heats of vaporization for non-associating organic liquids from the boiling points at various pressures. *Can. J. Chem.* **1986**, *64*, 635–640.

(31) Laato, H.; Lehtonen, P. Ueber die darstellung von Ameisensaeurephenylester und Phenyl-dichoromethylaether. *Suomen Kem. B* **1964**, *37*, 169–172.
(32) By A. N. Klym, in J.P.G.'s laboratory.
(33) The error estimates are based on statistical error propagation and take no account of possible systematic errors. Because the number of experiments is small, the true error is probably larger.
(34) Honig, B.; Nicholls, A. Classical electrostatics in biology and chemistry. *Science* **1995**, *268*, 1144–1149.
(35) Rashin, A. A.; Honig, B. Reevaluation of the Born model of ion hydration. *J. Phys. Chem.* **1985**, *89*, 5588–5593.
(36) Gilson, M.; Rashin, A.; Fine, R.; Honig, B. On the calculation of electrostatic interactions in proteins. *J. Mol. Biol.* **1985**, *184*, 503–516.
(37) Sharp, K.; Fine, R.; Honig, B. Computer simulations of the diffusion of a substrate to an active site of an enzyme. *Science* **1987**, *236*, 1460–1463.
(38) Sharp, K.; Honig, B. Lattice models of electrostatic interaction: the finite difference Poisson−Boltzmann method. *Chem. Scr.* **1989**, *29A*, 71–74.
(39) Nicholls, A.; Honig, B. A rapid finite difference alogrithm, utililizing successive over-relaxation to solve the Poisson−Boltzman equation. *J. Comput. Chem.* **1991**, *12*, 435–445.
(40) Rizzo, R. C.; Aynechi, T.; Case, D. A.; Kuntz, I. D. Estimation of absolute free energies of hydration using continuum methods: accuracy of partial charge models and optimization of nonpolar contributions. *J. Chem. Theor. Comput.* **2006**, *2*, 128–139.
(41) Bordner, A. J.; Cavasotto, C. N.; Abagyan, R. A. Accurate transferable model for water, n-octanol, and n-hexadecane solvation free energies. *J. Phys. Chem.* **2002**, *106*, 11009–11015.
(42) Sitkoff, D.; Sharp, K.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.
(43) PCModel, Serena Software, Bloomington, IN.
(44) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M., Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; 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, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*; Gaussian Inc.: Wallingford, CT, 2004.
(45) Nicholls, A. Unpublished data.
(46) Mobley, D. L.; Chodera, J. D.; Dill, K. A. Treating entropy and conformational changes in implicit solvent simulations of small molecules. Accepted, *J. Phys. Chem. B.*
(47) Grant, J. A.; Pickup, B. T.; Nicholls, A. A smooth permittivity function for Poisson−Boltzmann solvation methods. *J. Comput. Chem.* **2001**, *22*, 608–640.
(48) Rodinger, T.; Pomès, R. Enhancing the accuracy, the efficiency and the scope of free energy simulations. *Curr. Opin. Struct. Biol.* **2005**, *15*, 164–170.
(49) Shirts, M. R.; Mobley, D. L.; Chodera, J. D. Alchemical free energy calculations: ready for prime time? *Annu. Rep. Comput. Chem.* **2007**, *39*, 41–59.
(50) Shirts, M. R.; Pande, V. S. Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration. *J. Chem. Phys.* **2005**, *122*, 144107.
(51) Wu, D.; Kofke, D. A. Phase space overlap measures. I. Fail-safe bias detection in free energies calculated by molecular simulation. *J. Chem. Phys.* **2005**, *123*, 054103.
(52) Mobley, D. L.; Chodera, J. D.; Dill, K. A. On the use of orientational restraints and symmetry number corrections in alchemical free energy calculations. *J. Chem. Phys.* **2006**, *125*, 084902.
(53) Lu, N.; Adhikari, J.; Kofke, D. A. Variational formula for the free energy based on incomplete sampling in a molecular simulation. *Phys. Rev. E* **2003**, *68*, 026122.
(54) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod.* **2001**, *7*, 306–317.
(55) van der Spoel, D.; Lindahl, B.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: fast, flexible and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
(56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926.
(57) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general Amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
(58) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. Comparison of charge models for fixed-charge force fields: small-molecule hydration free energies in explicit solvent. *J. Phys. Chem. B* **2007**, *111*, 2242–2254.
(59) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J. Comput. Chem.* **2000**, *21*, 132–146.
(60) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641.
(61) We noticed that the AM1-BCC charges produced by ANTECHAMBER did not symmetrize the charges on all chemically equivalent atoms, as described in the original AM1-BCC paper of Jakalian et al.[59,60] The charges provided by C. Bayly came from a proprietary Merck-Frosst implementation, where charges on chemically equivalent atoms were in fact symmetrized.
(62) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Modell.* **2006**, *26*, 247–260.
(63) Bayly, C. I. Personal communication.
(64) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
(65) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268.
(66) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Phys. Rev. Lett.* **2001**, *91*, 140601.
(67) Lu, N. D.; Singh, J. K.; Kofke, D. A. Appropriate methods to combine forward and reverse free-energy perturbation averages. *J. Chem. Phys.* **2003**, *118*, 2977–2984.
(68) Sharp, K.; Jean-Charles, A.; Honig, B. A local dielectric constant model for solvation free energies which accounts for solute polarizability. *J. Phys. Chem.* **1992**, *96*, 3822–3828.
(69) Chipot, C. Rational determination of charge distributions for free energy calculations. *J. Comput. Chem.* **2003**, *24*, 409–415.
(70) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **2005**, *113*, 133–151.
(71) Rizzo, R. C.; Jorgensen, W. L. OPLS All-Atom Model for Amines: Resolution of the Amine Hydration Problem. *J. Am. Chem. Soc.* **1999**, *121*, 4827–4836.
(72) Schurhammer, R.; Engler, E.; Wipff, G. Hydrophobic ions in TIP5P water and at a water-chloroform interface: the effect of sign inversion investigated by MD and FEP simulations. *J. Phys. Chem. B* **2001**, *105*, 10700–10708.
(73) Chorny, I.; Dill, K. A; Jacobson, M. P. Surfaces affect ion pairing. *J. Phys. Chem.* **109**, *50*, 24056–24060.
(74) Pylasunov, A. V.; Shock, E. L. Thermodynamic functions of hydration of hydrocarbons at 298.15 K and 0.1 MPa. *Geochim. Cosmochim. Acta* **2000**, *64*, 439–468.
(75) Reported experimental uncertainties are often quite small or are difficult to extract. Several studies, however, have compared measurements of hydration free energies for the same compounds from different groups and found variations in the range of 0.01−0.1 kcal/mol.[2,74] Uncertainty estimates for some common compounds have been suggested to be around 0.2 kcal/mol but sometimes larger.[76] Therefore, here, we take 0.2 kcal/mol as a (hopefully conservative) estimate.
(76) Abraham, M. H.; Whiting, G. S.; Fuchs, R.; Chambers, E. J. Thermodynamics of solute transfer from water to hexadecane. *J. Chem. Soc., Perkin Trans. 2* **1990**, 291–300.
(77) Guthrie, J. P.; Pike, D. C.; Lee, Y. C. Equilibrium constants and heats of formation of methyl esters and *N*,*N*-dimethyl amides of substituted benzoic acids. *Can. J. Chem.* **1992**, *70*, 1671–1683.
(78) Przyjazny, A.; Janicki, W.; Chrzanowsi, W.; Staszewski, R. Headspace gas chromatographic determination of distribution coefficients of selected organosulphur compounds and their dependence on some parameters. *J. Chromatogr.* **1983**, *280*, 249–260.
(79) Guthrie, J. P. Unpublished, based on older vapor pressure data.
(80) Calculated from solubility and vapor pressure: see Methods.
(81) Wolfenden, R.; Liang, Y.-L.; Matthews, M.; Williams, R. Cooperativity and anticooperativity in solvation by water: imidazoles, quinones, nitrophenols, nitrophenolate, and nitrothiophenolate ions. *J. Am. Chem. Soc.* **1987**, *109*, 463–466.

JM070549+