Modeling aqueous solvation with semi-explicit assembly

Christopher J. Fennell[a], Charles W. Kehoe[b], and Ken A. Dill[a,1]

[a]Department of Pharmaceutical Chemistry, and [b]Graduate Group in Bioinformatics, University of California, San Francisco, CA 94143

Contributed by Ken A. Dill, December 1, 2010 (sent for review October 1, 2010)

We describe a computational solvation model called semi-explicit assembly (SEA). SEA water captures much of the physics of explicit-solvent models but with computational speeds approaching those of implicit-solvent models. We use an explicit-water model to precompute properties of water solvation shells around simple spheres, then assemble a solute's solvation shell by combining the shells of these spheres. SEA improves upon implicit-solvent models of solvation free energies by accounting for local solute curvature, accounting for near-neighbor nonadditivities, and treating water's dipole as being asymmetrical with respect to positive or negative solute charges. SEA does not involve parameter fitting, because parameters come from the given underlying explicit-solvation model. SEA is about as accurate as explicit simulations as shown by comparisons against four different homologous alkyl series, a set of 504 varied solutes, solutes taken retrospectively from two solvation-prediction events, and a hypothetical polar-solute series, and SEA is about 100-fold faster than Poisson–Boltzmann calculations.

free energy | implicit solvent | transfer

We describe here an approach for computing the free energies of solvation of solutes in water. Aqueous solvation has been modeled at different levels, ranging from detailed quantum mechanics simulations of few-molecule clusters (1, 2), to faster classical simulations using up to tens of thousands of explicit molecules (3–10), to very fast models in which water is treated implicitly as a simple uniform continuous medium (11–17). For large computations, such as those in typical biomolecule simulations, explicit-water modeling can be slow and expensive, so it is common to use implicit water instead. However, implicit models often require trade-offs in the physics that can limit their accuracies. For example, water is typically treated as a continuum rather than individual particles, and this neglects discrete microscopic effects; nonpolar solvation effects are often assumed to depend only on surface area A (expressed as γA), and not on detailed dispersive interactions and collective consequences of solute shape (18–20).

It would be useful to have a computational model of water that is both fast—approaching the speeds of the fastest implicit-solvent models—and that captures the physics and the transferability of explicit-solvent models. Toward this goal, various improvements of implicit models have been introduced (21, 22), explicit solvents have been coarse-grained (23, 24), and hybrid explicit-implicit models have been developed (25–29). Here, we take a different approach. We precompute solvation properties of water in explicit-solvent simulations of simple spheres, which we then apply in summations over assemblies about arbitrary solutes. As the details of the solvation response come entirely from the physics of an explicit solvent, this model lacks free parameters from statistical fits to solute molecular transfer free energies, resulting in a wide transferability. We call this approach semi-explicit assembly, or SEA water.

Theory

In the SEA approach, we assume that the free energy of interaction, ΔG, of a solute molecule with its solvating water molecules, is a sum of three terms,

$$\Delta G = \Delta G_{\mathrm{np}} + \Delta G_{\mathrm{pol,surf}} + \Delta G_{\mathrm{pol,bulk}}. \quad [1]$$

ΔG_{np} is the free energy of forming the cavity the solute occupies in water and includes the dispersion interactions between the solute and the waters. For purely nonpolar solutes, such as simple hydrocarbons, ΔG_{np} is the only substantial contribution to ΔG. For solute molecules that are polar or charged, the total solvation free energy also has two additional terms (see Eq. **1**): $\Delta G_{\mathrm{pol,surf}}$ describes the electrostatic interactions of the solute with the immediate-neighbor first-shell water molecules that surround it, and $\Delta G_{\mathrm{pol,bulk}}$ describes the electrostatic interactions of the solute with the water molecules that are more distant than the first solvation shell.

For the nonpolar component ΔG_{np}, the SEA approach is described in detail elsewhere (20). In the present paper, we describe the SEA treatment of the surface and bulk polar terms. We compute these polar components using two steps: (*i*) a slow series of presimulations for a given solvent and (*ii*) a fast summation of component free energies for a given solute.

SEA Step 1: Precalculations on Solute Spheres in Explicit Water. SEA treats a solute molecule as a collection of "atomic" spheres of different types and sizes (i.e., different Lennard–Jones parameters and different partial charges). The first step in SEA modeling is to precalculate the positions and orientations of explicit-water molecules around those component building-block spheres. SEA can use any explicit-solvent model. Here, to illustrate the principle, we use the TIP3P water model (30).

We first solvate a series of individual spheres in baths of explicit water. Spheres span a range of different sizes and Lennard–Jones parameters (6ϵ values and 6 σ values) and a range of 11 different partial charges. In total, we do precomputations on 396 different types of spheres in TIP3P water. Each sphere is simulated for 10 ns at 1 atm and 300 K. From each such simulation, we harvest the statistical properties of water distances and orientations around the sphere. For each type of sphere, we calculate the average distance between the solute and the water (the peak of the solute to water radial distribution function; see Table S1). Calculating radial distribution functions for water around these spheres, we found that the width of the first hydration shell extends out 1 to 3 Å from the van der Waals surface of the spheres, depending on the solute charge. We use a 2-Å cutoff distance from the surface of each solute to define the "first shell" (Fig. 1*A*) in order to exclude second hydration shell water population, which begins to grow beyond this distance around the more highly charged solutes. Within this region, we calculate the average

Author contributions: C.J.F., C.W.K., and K.A.D. designed research; C.J.F. and C.W.K. performed research; C.J.F. and C.W.K. contributed new reagents/analytic tools; C.J.F., C.W.K., and K.A.D. analyzed data; and C.J.F., C.W.K., and K.A.D. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: dill@maxwell.ucsf.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1017130108/-/DCSupplemental.

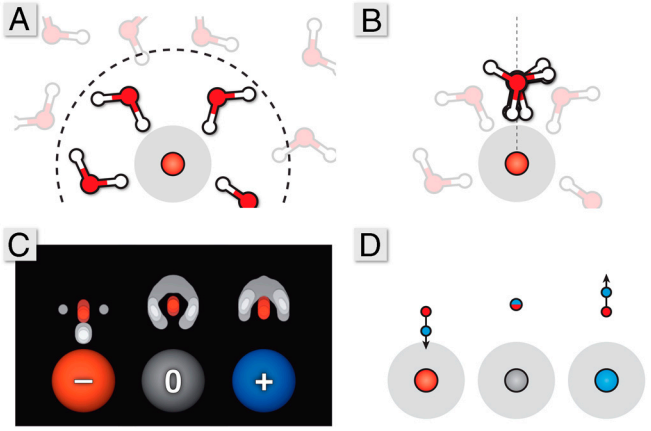

Fig. 1. The precomputation step. (*A*) Simulate the waters around sphere. (*B*) Superimpose all the first-shell waters onto a common axis relative to the solute–sphere center. (*C*) Create solvent atomic density maps. (*D*) Convert the density maps into dipole representations.

number of water molecules and generate nearest-neighbor distribution functions. These distribution functions tell us the average distance between first-shell waters about a given solute sphere (see *SI Text*, Fig. S1, and Table S2).

Around charged solute spheres, water in the first solvation shell will adopt a distribution of orientations and distances with respect to the solute sphere centers. To determine an average representation of a surface water, we align all of the first-shell water molecules onto a common axis normal to the surface of the solute sphere (Fig. 1*B*). The solvent atom locations are binned about this axis to form density maps of the surface water charges (Fig. 1*C*). From these maps, we extract dipole moments that capture the first-shell polar response of these local water molecules (Fig. 1*D*). In addition to a dipolar representation, an alternative quadrupole moment can be extracted (see *SI Text* and Fig. S2). Not unexpectedly, these water dipoles orient most strongly around spheres having the strongest electrostatic potentials, i.e., around spheres having the largest absolute charges.

The average first-shell water dipole moment has a simple functional relationship vs. the strength of the electrostatic field from the solute sphere (see *SI Text* and Figs. S3–S5). This observed relationship allows us to replace a simulation lookup with a sigmoidal function,

$$f(x) = \frac{1}{c_0 + \exp(c_1 \cdot x)} + c_2, \qquad [2]$$

where the coefficients c_0, c_1, and c_2 come from curve fits of the simulations. Although any saturating function, like the Langevin equation (22), would likely suffice in this fit to the explicit-solvent response, Eq. **2** is flexible, allowing it to retain subtle microscopic effects like biased water orientations around neutral solutes (31). Water's dipole is asymmetric; hence there is an asymmetry in the resulting function, with separate fits needed for the positive and negative field cases (32–35). Around positive spheres, the limit of the sigmoidal curve coincides with the dipole moment of the TIP3P water model because water's average dipole points away from the solute surface along the normal vector. Around negative spheres, water points one hydrogen toward the solute center (Fig. 1*C*) and only a partial projection of a water molecule's dipole is along this normal vector.

SEA Step 2: Assembling the Solvation Response about an Arbitrary Molecule. Once the precomputations have been performed, the SEA model can rapidly calculate the hydration free energy of a given solute. The nonpolar component is accumulated using our previously described approach (20), whereas the polar component is the sum of surface and bulk terms. To determine the surface polarization term, $\Delta G_{\text{pol,surf}}$, surface waters must be placed in physically representative locations and orientations. We generate a solvent-accessible dot surface about the solute of interest (Fig. 2). Rather than rolling a hard sphere of some fixed size over the molecule to generate a Lee–Richards or Connolly surface (36, 37), we use instead our precomputed average separations between a water molecule and the sphere center (20). This strategy captures various physical aspects of solvation, such as the fact that water molecules are held more tightly to solutes having stronger solute–solvent interactions.

We select a solvation site at random on this surface and compute the electric field at this point due to all (N) small-molecule solute atoms using

$$\mathbf{E} = \sum_{i=1}^{N} \frac{1}{4\pi\varepsilon_0} \frac{q_i}{r_{ij}^2} \hat{\mathbf{r}}_{ij}, \qquad [3]$$

where q_i is the charge on solute atom i and r_{ij} is the distance between site j and solute atom i. Using the electric field association functions (Eq. **2**), the appropriate dipole (or quadrupole) is placed at this site along the electric field line. After placement, surface water sites that are closer than the distance to the peak of this dipole's nearest-neighbor distribution are eliminated. The random placement procedure then continues until all possible solvation sites are occupied or eliminated.

After generating a solvent configuration around a solute, the surface interactions are accumulated into $\Delta G_{\text{pol,surf}}$ using

$$\Delta G_{\text{pol,surf}} = \Delta G_{\text{sol-solv}} + \Delta G_{\text{solv-solv}}, \qquad [4]$$

where sol refers to solute particles and solv refers to solvent particles. For the small-molecule solutes investigated here, $\Delta G_{\text{sol-solv}}$ is accumulated from a pairwise Coulombic sum between all solute atoms and all water site partial charges, whereas $\Delta G_{\text{solv-solv}}$ is a pairwise Coulombic sum between all partial charges of differing water sites. Because positional and orientational distributions of surface water molecules are factored into the dipole representations, this pairwise sum approximates a free energy rather than simple potential energy. It should be noted that although these simple charge–dipole and dipole–dipole sums approximate the enthalpic enhancement and entropic loss from solvent electrostriction, a more detailed accounting of the solvation specifics could lead to future improvements in the technique.

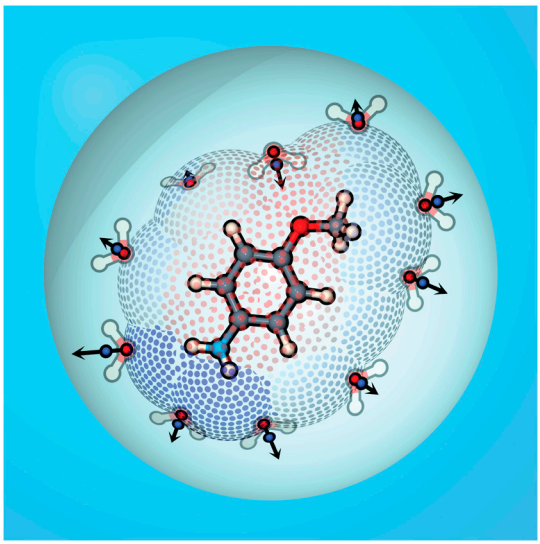

Fig. 2. An illustration of SEA sampling process around p-methoxyaniline. Semi-explicit dipoles are placed along the solvent-accessible dot surface according to the local electric field within a continuum dielectric cavity.

To estimate the total polar component of solvation, we need the first-shell/surface component described above and the bulk component of the electrostatic free energy of solvation. We estimate this quantity using the Onsager reaction field (38),

$$\Delta G_{\text{pol,bulk}} = \frac{\varepsilon - 1}{2\varepsilon + 1} \cdot \frac{\boldsymbol{\mu}^2}{r_c^3}, \quad [5]$$

where $\boldsymbol{\mu}$ is the dipole vector of the spherical solute cavity of radius r_c and ε is the dielectric constant of the external continuum. Here, r_c extends from the geometric center of the solute out to the outer boundary of the farthest semi-explicit shell (see the spherical cavity in Fig. 2), and the cavity dipole is accumulated from a simple sum over all point charges within r_c, which includes both the solute and the semi-explicit partial charge sites. This type of reaction field works well for neutral polar molecules, which are our primary focus here, but for solutes having a formal net charge, an additional component will be needed to account for Born effects (25, 39, 40).

To get proper averages, we repeat the random placement of the first solvation shell a set number of times to estimate the variance of the ΔG_{pol} as a function of water placement. This variance tells us how many additional sampling iterations are needed to converge an averaged ΔG to a desired standard error. We then Boltzmann-weight these solvation-shell snapshots and calculate the probability, p_i, of shell configuration i using (41)

$$p_i = \frac{e^{-\Delta G_i / k_B T}}{\sum_{j=1}^{N} e^{-\Delta G_j / k_B T}}, \quad [6]$$

where N is the total number of solvation configuration samples. The average solvation free energy is the weighted sum over the different shell configurations,

$$\langle \Delta G \rangle = \sum_{i=1}^{N} \Delta G_i \cdot p_i. \quad [7]$$

This Boltzmann-weighted sum enhances contribution from more probable, and correspondingly deemphasizes less probable, solvation configurations.

Results and Discussion

We now compare the SEA water model to experiments, to explicit-solvent simulations, and to Poisson–Boltzmann (PB) and generalized Born (GB) models of solvation.

Testing SEA on Homologous Series of Alkyl Chains Terminated by Different Functional Groups. Fig. 3 shows calculated solvation free energies, ΔG, for four series of alkyl chains terminated by different organic functional groups. We compare SEA to experiments, to TIP3P simulations, and to GB. First, SEA gives good agreement with the TIP3P results, within 1 kcal/mol in all cases. Because the parameters used in SEA water modeling are identical to those of TIP3P, this shows that SEA's sampling and regional additivity approximation are not degrading the quality of predictions relative to the much more expensive explicit-solvent simulations.

Second, although TIP3P and SEA both give good estimates of experimental ΔG values for acetates and phenyl groups, both methods predict solvation free energies for the alcohols and amines that are too positive. This indicates that the combination of General Amber Force Field (GAFF) and the AM1-BCC

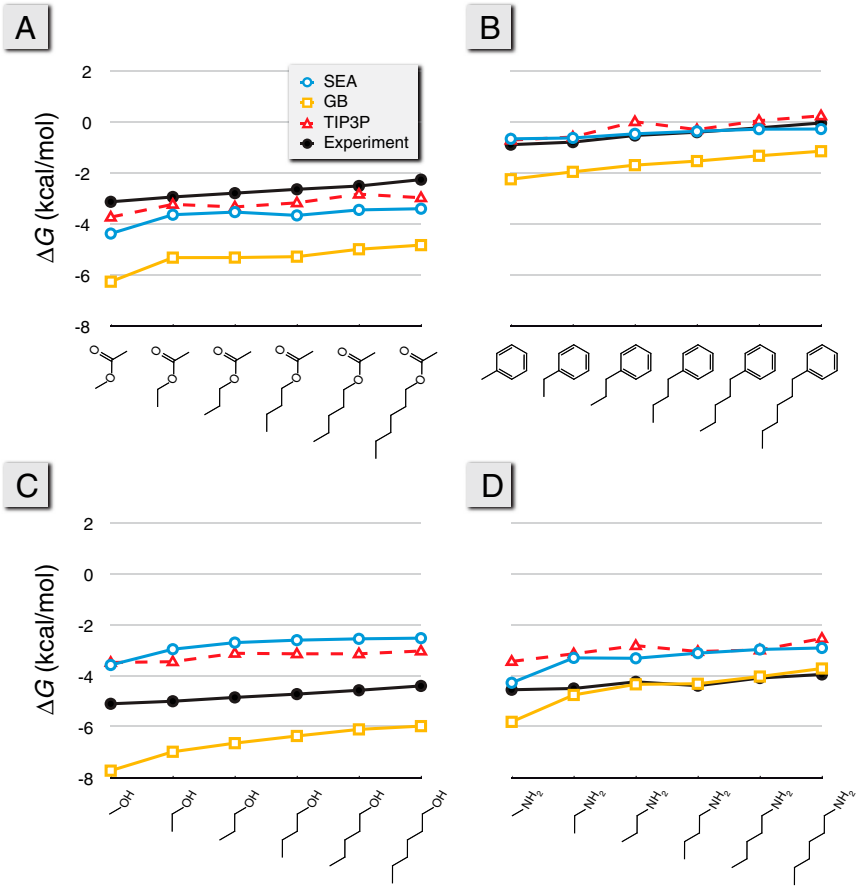

Fig. 3. Functional group comparisons for the linear alkyl series of (*A*) acetates, (*B*) phenyls, (*C*) alcohols, and (*D*) amines. Experimental results come from ref. 55 and the TIP3P results from ref. 10. The GB results are from Amber 10 using iGB5 and γA with $\gamma = 5$ cal mol^{-1} Å^{-2} (13).

partial charges used in the underlying explicit-solvation model needs improvement for these functional groups. Such errors have been noted before in the explicit-solvent models (4, 6, 10).

Third, we find that GB gives free energies of solvation that are too negative in all cases except for the amines, where it matches experiments quite accurately. The errors in GB can be partly corrected by using the nonpolar SEA term in place of the standard γA term, resulting in a general 0.2–0.4 kcal/mol improvement for these molecules.

Testing SEA on a Diverse Set of 504 Small Molecules. Next, we explore a diverse solute set, containing 504 different small molecules (10, 17, 42). It is useful because of the availability of both experimental data and of extensive TIP3P explicit-solvent simulations (10).

Fig. 4 compares GB, PB, and SEA with the TIP3P simulation results. Taking TIP3P as the "gold standard" for this test, Fig. 4 shows systematic improvement from GB to PB to SEA. SEA is within 1 kcal/mol root mean square error (RMSE) relative to TIP3P. Thus, although there is a large gain in computational efficiency from TIP3P explicit-solvent simulations to SEA, there is little loss in predictive accuracy for computing solvation free energies.

The PB and GB implicit-solvent models tend to solvate many of these molecules more favorably than SEA or explicit solvent. There are some exceptions where GB returns significantly less favorable solvation free energies than TIP3P, and these are nearly all molecules with buried amines, like piperazine and triethylamine. The balance of Born radii about amines has been recognized as a reason behind overly strong salt bridges in protein simulations using GB, and suggested corrections have included resizing the radii of atoms attached to these nitrogens to fit to explicit-solvent energetics (43). Both PB and SEA appear to avoid suffering from this issue.

Speed vs. Accuracy of SEA. Fig. 5 compares the performance (computational speed per ligand vs. the RMSE free-energy accuracy to experiment) of the various computational methods applied to the test set of 504 small molecules described above. Although there are many different implicit-solvent implementations available, our test here is intended only as a general sketch, not an exhaustive and detailed comparison. Here is the range: TIP3P is about 10^8-fold slower than GB (the fastest method) in calculating solvation free energies. And, RMSE errors are about 2.8 kcal/mol in GB or 1.2 kcal/mol in TIP3P. SEA provides a good compromise. In SEA, RMSE errors are about 1.3 kcal/mol, but it is only 10-fold slower than GB*.

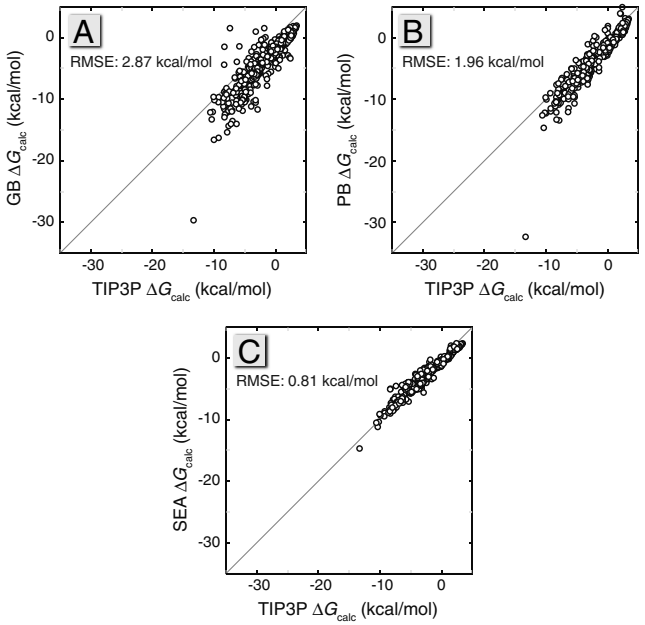

Fig. 4. Scatter plots of (*A*) GB (13), (*B*) PB (53), and (*C*) SEA air–water ΔG transfer values against TIP3P calculations (10). If the single largest outlier (triacetylglycerol) in the GB and PB calculations is removed, the RMSEs decrease to 2.78 and 1.77 kcal/mol, respectively.

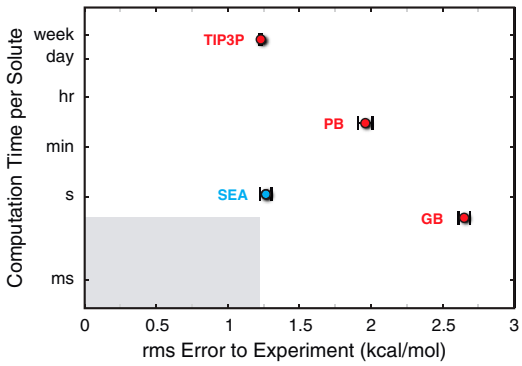

Fig. 5. The average execution time per solute vs. the RMSE to experiment for calculations on the 504 small-molecule solute set. The vertical axis is a logarithmic scale. The light-gray region indicates the region of performance faster than GB and more accurate than TIP3P, a target realm for future solvent models.

Retrospective Analysis of Statistical Assessment of the Modeling of Proteins and Ligands (SAMPL) Blind-Test Solvation Predictions. Recently, a community-wide blind prediction event for small-molecule solvation free energies, called SAMPL, has been run by OpenEye Software (44). The SEA method that we describe in this paper has not yet been tested in that event. (We tested an incomplete version without water dipoles or solute conformational sampling in SAMPL2.) Here, as a second-best alternative to a blind test, we retrospectively test the present SEA approach predictions on the solute molecules from both prior SAMPL events.

As we found in the other two comparisons above, the SEA model performs at about the same level of accuracy as the TIP3P model on which our model is based. Fig. 6*A* shows the calculated ΔG vs. experiment for the 56 SAMPL1 molecules. The resulting scatter is similar to that observed from explicit-solvent simulations (45). The RMSE of SEA vs. experiments is 4.1 kcal/mol, compared to 3.8 kcal/mol for TIP3P. Both SEA and explicit solvent have the same problems: They do not handle well molecules containing sulfur or phosphorus, a difficulty that is likely due to the GAFF force field parameters for these atom types (45).

Fig. 6*B* shows 23 SAMPL2 molecules, and the RMSE for this small set is under 2 kcal/mol. The two largest outliers are *d* xylose and *d* glucose. For these sugars, SEA predicts solvation free energies that are 4 and 5 kcal/mol less favorable than experiments. Taken together with the alcohol comparisons in Fig. 3*D*, it suggests a problem in the chosen force field parameters for hydroxyl groups. Although it would be unfair to compare our current retrospective results with the prospective tests made by other groups in SAMPL, nevertheless the present tests indicate that SEA is comparable to the best current prediction methods. Because the SEA method does not involve the fitting of free parameters, it is reasonable to expect that SEA should perform with about the same RMSE accuracy as TIP3P simulations in such future tests.

*Note, however, for speed comparisons that explicit simulations can automatically include multiple solute conformers, whereas other methods' speed evaluations are per set of 25 conformers.

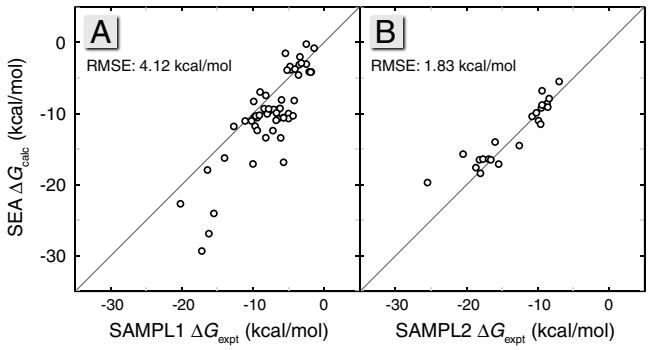

Fig. 6. The calculated ΔG from SEA vs. experimental values for the (*A*) 56 blind prediction molecules of the SAMPL1 event and the (*B*) 23 blind prediction molecules in the SAMPL2 event.

SEA Fixes Some Physical Flaws in Implicit-Solvent Models. Because water's dipole is asymmetric, it means that positively charged atoms of a given radius solvate less favorably than negatively charged atoms of the same radius (32–35, 46). Such asymmetries in the physics are not directly captured in implicit-solvent models. In implicit-solvent models, this asymmetry is handled instead by empirical readjustment of solvation radii. To understand this asymmetry more quantitatively, Mobley et al. created a series of polar molecules that they solvated in explicit solvent (35). Their fictitious benzene-like "bracelet" molecules were created to have either positive heads or negative heads and were otherwise net neutral. The difference in free energy of solvation of the positive-head bracelet and negative-head bracelet was found to be about 10 kcal/mol. Zero difference in free energy would have been expected from implicit-solvent models. This is a useful test set of molecules for exploring whether SEA water modeling is able to capture the asymmetry in solvation with solute charge.

Fig. 7 shows the results for bracelets of different size. Although SEA does capture the physics of the asymmetry, it underestimates the free-energy difference of the asymmetry. We find, however, that if we go beyond just using dipolar representation of water in SEA modeling and use a quadrupolar representation instead, the agreement with TIP3P simulations is improved. A quadrupole gives a better fit to the water response map around atoms having a formal negative charge, like the head atom on the negative bracelets. Interestingly, using the quadrupolar representation in SEA water on the 504 small-molecule set leads to no collective decrease in rms error with the experimental or the TIP3P results. This indicates that although the quadrupolar representation improves the calculation accuracy in select cases, the simpler and more computationally efficient dipolar representation captures much of the physics necessary for general applications.

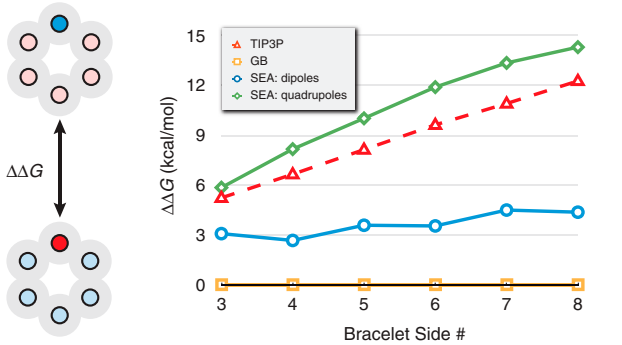

Fig. 7. The solvation asymmetry ($\Delta\Delta G$) upon partial charge inversion for model polygons studied in ref. 35. In this case, one atom of each polygon has a formal charge, whereas the neutralizing counter charge is distributed evenly among the remaining atoms, as shown in the six-sided bracelet illustration on the left.

Conclusions

We have described a method, called SEA, for computing the free energies of solvation of solutes in water. SEA uses precomputations with some chosen explicit-solvation model and then assembles solvation shells of water in order to compute solvation free energies of arbitrary solutes. Here, we performed these precomputations using TIP3P water, though more complex models, such as polarizable water models, or nonaqueous solvents would also be practical. SEA aims to correct some of the physical deficits of traditional implicit-solvent models, such as water's asymmetrical dipole, the lack of particulate water in the first solvation shell, and shape and microscopic interaction effects in nonpolar cavity formation. We have shown that SEA water predicts the solvation free energies for a range of nonpolar and polar small molecules at approximately the same accuracy as in TIP3P computer simulations, but it runs much faster than explicit simulations, much closer to the speeds of implicit-solvent models.

Computational Methods

Explicit Simulations. The simulations of model spheres were performed using GROMACS 4.0.4 (47, 48). In these calculations, statistics for four spheres were accumulated simultaneously by bonding them in a tetrahedral geometry with 20-Å edge lengths between each particle. Partial charge values of these sites ranged from -1 to $+1$ in increments of 0.2, and they were assigned to each tetrahedron in a manner that kept the explicit simulation net neutral. Separate simulations were performed for different sets of Lennard–Jones parameters. Model sphere diameters included $\sigma = 1.4$, 2.2, 3.0, 3.8, 4.6, and 5.4 Å, whereas well depths included $\epsilon = 0.015625$, 0.03125, 0.0625, 0.125, 0.25, and 0.5 kcal/mol. Each of the resulting 108 tetrahedral arrangements was solvated in a cubic box of TIP3P water with a 12-Å buffer of water between any solute particle and face of the simulation cell.

Each solvated tetrahedron was simulated under periodic boundary conditions in the isothermal-isobaric ensemble at 300 K and 1 atm using the Nose–Hoover thermostat and Parrinello–Rahman barostat with time constants of 1 and 10 ps, respectively. The dynamics were propagated with the leapfrog integrator with a timestep of 2 fs. The Lennard–Jones interactions were switched off between 8 and 9 Å, with long-ranged energy and pressure terms included. Smooth particle mesh Ewald (49) was used for electrostatics with a real-space cutoff of 10 Å, spline order of 4, Fourier spacing of 1.2 Å, and relative energy tolerance of 10^{-5}. Trajectory snapshots were saved at 1-ps intervals for the postprocessing described in the main text.

To investigate the possible effects of the tetrahedral setup geometry, we tested cubic, linear, and even net-charged single point arrangements and found no statistically significant difference in the resulting solvent quantities.

Details of the Implementation. Solvent-accessible surfaces for SEA calculations were constructed as described in our previous work (20). The surface detail was set to approximately 80 surface points per solute atom. The dielectric constant of the bulk surroundings was set to 78 for the reaction field term. A minimum of five semi-explicit configurations were sampled, and this iteration process was continued until the accumulated ΔG_{pol} converged with an error of ~0.05 kcal/mol.

Force Fields And Conformation Sampling. The GAFF (50) was used to assign Lennard–Jones parameters for all of the small molecules. AM1-BCC partial charges (51) were assigned to initial minimized structures using the ANTECHAMBER program in AMBER (52). As these structures may not be the ideal representative structures of the molecules in solution, we generated short conformation ensembles for each from 500 ps of MD in explicit solvent using GROMACS 4.0 (47, 48). The results were clustered

to obtain an average representation of the small molecule, and the partial charges were reassigned using this averaged conformation.

In the cases where conformation plays a role in the calculated transfer free energies, an ensemble of conformations is necessary to obtain an accurate estimation. We generated solute conformation ensembles from 5-ns explicit-solvent MD simulations with 2-, 20-, and 200-ps snapshot intervals for all small molecules investigated in this study. The various conformation sets all gave results that overlapped within error, so we used the ensembles generated from the MD simulations with 200-ps snapshot intervals (25 conformations) for maximum efficiency. The free energies reported for each solute from SEA and the implicit-solvent models come from averages over these 25 conformations.

Implicit-Solvent Calculations. Implicit-solvent results were obtained for the PB and GB methods using the adaptive Poisson-Boltzmann solver (APBS) v 1.0 (53) and the Onufriev–Bashford–Case implementation of GB in AMBER 10 (13, 54) (igb = 5), respectively. As with the SEA and TIP3P calculations, atom Lennard–Jones parameters were assigned using GAFF with AM1-BCC partial charges. For the PB calculations, the nonpolar term was computed using the APBS dispersion incorporation method (18). The polar term was computed using the van der Waals surface as the dielectric boundary, and the PB equation was solved on a $65 \times 65 \times 65$ grid with 0.25-Å spacing. In the GB calculations, γA was used for the nonpolar term with a γ of 5 cal mol^{-1} Å^{-2}. Both the PB and GB calculations used a continuum dielectric of 78.

ACKNOWLEDGMENTS. The authors thank David Mobley (University of New Orleans) and Jordi Villà i Frexia (Parc de Recerca Biomèdica de Barcelona) for helpful discussions. The authors also appreciate the financial support provided by NIH Grant GM63592.

1. Silvestrelli PL, Parrinello M (1999) Structural, electronic, and bonding properties of liquid water from first principles. *J Chem Phys* 111:3572–3580.
2. Sharma M, Resta R, Car R (2007) Dipolar correlations and the dielectric permittivity of water. *Phys Rev Lett* 98:247401.
3. Levy RM, Gallicchio E (1998) Computer simulations with explicit solvent: Recent progress in the thermodynamic decomposition of free energies and in modeling electrostatic effects. *Annu Rev Phys Chem* 49:531–567.
4. Villa A, Mark AE (2002) Calculation of the free energy of solvation for neutral analogs of amino acid side chains. *J Comput Chem* 23:548–553.
5. Shirts MR, Pitera JW, Swope WC, Pande VS (2003) Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *J Chem Phys* 119:5740–5761.
6. MacCallum JL, Tieleman DP (2003) Calculation of the water-cyclohexane transfer free energies of neutral amino acid side-chain analogs using the opls all-atom force field. *J Comput Chem* 24:1930–1935.
7. Deng Y, Roux B (2004) 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* 108:16567–16576.
8. Hess B, van der Vegt NFA (2006) Hydration thermodynamic properties of amino acid analogues: A systematic comparison of biomolecular force fields and water models. *J Phys Chem B* 110:17616–17626.
9. Mobley DL, Dumont E, Chodera JD, Dill KA (2007) Comparison of charge models for fixed-charge force fields: Small-molecule hydration free energies in explicit solvent. *J Phys Chem B* 111:2242–2254.
10. Mobley DL, Bayly CI, Cooper MD, Shirts MR, Dill KA (2009) Small molecule hydration free energies in explicit solvent: An extensive test of fixed-charge atomistic simulations. *J Chem Theory Comput* 5:350–358.
11. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *J Am Chem Soc* 112:6127–6129.
12. Cramer CJ, Truhlar DG (1992) An SCF solvation model for the hydrophobic effect and absolute free energies of aqueous solvation. *Science* 256:213–217.
13. Onufriev A, Bashford D, Case D (2000) Modification of the generalized born model suitable for macromolecules. *J Phys Chem B* 104:3712–3720.
14. Luo R, David L, Gilson MK (2002) Accelerated Poisson-Boltzmann calculations for static and dynamic systems. *J Comput Chem* 23:1244–1253.
15. Gallicchio E, Levy RM (2004) AGBNP: An analytic implicit solvent model suitable for molecular dynamics simulations and high-resolution modeling. *J Comput Chem* 25:479–499.
16. Tomasi J, Mennucci B, Cammi R (2005) Quantum mechanical continuum solvation models. *Chem Rev* 105:2999–3093.
17. Rizzo RC, Aynechi T, Case DA, Kuntz ID (2006) Estimation of absolute free energies of hydration using continuum methods: Accuracy of partial charge models and optimization of nonpolar contributions. *J Chem Theory Comput* 2:128–139.
18. Wagoner JA, Baker NA (2006) Assessing implicit models for nonpolar mean solvation forces: The importance of dispersion and volume terms. *Proc Natl Acad Sci USA* 103:8331–8336.
19. Tan C, Tan Y-H, Luo R (2007) Implicit nonpolar solvent models. *J Phys Chem B* 111:12263–12274.
20. Fennell CJ, Kehoe C, Dill KA (2010) Oil/water transfer is partly driven by molecular shape, not just size. *J Am Chem Soc* 132:234–240.
21. Papazyan A, Warshel A (1998) Effect of solvent discreteness on solvation. *J Phys Chem B* 102:5348–5357.
22. Florián J, Warshel A (1999) Calculations of hydration entropies of hydrophobic, polar, and ionic solutes in the framework of the Langevin dipoles solvation model. *J Phys Chem B* 103:10282–10288.
23. Marrink SJ, de Vries AH, Mark AE (2004) Coarse grained model for semiquantitative lipid simulations. *J Phys Chem B* 108:750–760.
24. Izvekov S, Voth GA (2006) Multiscale coarse-graining of mixed phospholipid/cholesterol bilayers. *J Chem Theory Comput* 2:637–648.
25. King G, Warshel A (1989) A surface constrained allatom solvent model for effective simulations of polar solutions. *J Chem Phys* 91:3647–3661.
26. Beglov D, Roux B (1994) Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations. *J Chem Phys* 100:9050–9063.
27. Lounnas V, Lüdemann SK, Wade RC (1999) Towards molecular dynamics simulation of large proteins with a hydration shell at constant pressure. *Biophys Chem* 78:157–182.
28. Im W, Bernèche S, Roux B (2001) Generalized solvent boundary potential for computer simulations. *J Chem Phys* 114:2924–2937.
29. Lee MS, Salsbury FR, Jr, Olson MA (2004) An efficient hybrid explicit/implicit solvent method for biomolecular simulations. *J Comput Chem* 25:1967–1978.
30. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935.
31. Ashbaugh HS (2000) Convergence of molecular and macroscopic continuum descriptions of ion hydration. *J Phys Chem B* 104:7235–7238.
32. Latimer WM, Pitzer KS, Slansky CM (1939) The free energy of hydration of gaseous ions, and the absolute potential of the normal calomel electrode. *J Chem Phys* 7:108–111.
33. Hummer G, Pratt LR, García AE (1996) Free energy of ionic hydration. *J Phys Chem* 100:1206–1215.
34. Rajamani S, Ghosh T, Garde S (2004) Size dependent ion hydration, its asymmetry, and convergence to macroscopic behavior. *J Chem Phys* 120:4457–4466.
35. Mobley DL, Barber AE, II, Fennell CJ, Dill KA (2008) Charge asymmetries in hydration of polar solutes. *J Phys Chem B* 112:2405–2414.
36. Lee B, Richards RM (1971) The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol* 55:379–400.
37. Connolly ML (1983) Analytical molecular surface calculation. *J Appl Cryst* 16:548–558.
38. Onsager L (1936) Electric moments of molecules in liquids. *J Am Chem Soc* 58:1486–1493.
39. Åqvist J (1990) Ion-water interaction potentials derived from free energy perturbation slmulations. *J Phys Chem* 94:8021–8024.
40. Darden T, Pearlman D, Pedersen LG (1998) Ionic charging free energies: Spherical versus periodic boundary conditions. *J Chem Phys* 109:10921–10935.
41. Fennen J, Torda AE, van Gunsteren WF (1995) Structure refinement with molecular dynamics and a boltzmann-weighted ensemble. *J Biomol NMR* 6:163–170.
42. Bordner AJ, Cavasotto CN, Abagyan RA (2002) Accurate transferable model for water, n-octanol, and n-hexadecane solvation free energies. *J Phys Chem B* 106:11009–11015.
43. Geney R, Layten M, Gomperts R, Hornak V, Simmerling C (2006) Investigation of salt bridge stability in a generalized born solvent model. *J Chem Theory Comput* 2:115–127.
44. Guthrie JP (2009) A blind challenge for computational solvation free energies: Introduction and overview. *J Phys Chem B* 113:4501–4507.
45. Mobley DL, Bayly CI, Cooper MD, Dill KA (2009) Predictions of hydration free energies from all-atom molecular dynamics simulations. *J Phys Chem B* 113:4533–4537.
46. Purisima EO, Sulea T (2009) Restoring charge asymmetry in continuum electrostatics calculations of hydration free energies. *J Phys Chem B* 113:8206–8209.
47. Berendsen HJC, van der Spoel D, van Drunen R (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Comput Phys Commun* 91:43–56.
48. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4:435–447.
49. Essman U, et al. (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103:8577–8592.
50. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25:1157–1174.
51. Jakalian A, Bush BL, Jack DB, Bayly CI (2000) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J Comput Chem* 21:132–146.
52. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* 25:247–260.
53. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98:10037–10041.
54. Case DA, et al. (2008) AMBER 10. (University of California, San Francisco).
55. Abraham MH, Whiting GS, Fuchs R, Chambers EJ (1990) Thermodynamics of solute transfer from water to hexadecane. *J Chem Soc Perk T 2* 291–300.