JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

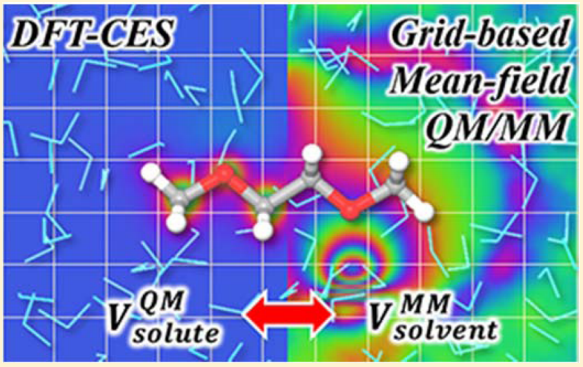
A Seamless Grid-Based Interface for Mean-Field QM/MM Coupled with Efficient Solvation Free Energy Calculations

Hyung-Kyu Lim,[§] Hankyul Lee,[§] and Hyungjun Kim*

Graduate School of Energy, Environment, Water, and Sustainability (EEWS), Korea Advanced Institute of Science and Technology (KAIST), Yuseong-gu, Daejeon 305-701, Korea

Supporting Information

ABSTRACT: Among various models that incorporate solvation effects into first-principles-based electronic structure theory such as density functional theory (DFT), the average solvent electrostatic potential/molecular dynamics (ASEP/MD) method is particularly advantageous. This method explicitly includes the nature of complicated solvent structures that is absent in implicit solvation methods. Because the ASEP/MD method treats only solvent molecule dynamics, it requires less computational cost than the conventional quantum mechanics/molecular mechanics (QM/MM) approaches. Herein, we present a real-space rectangular grid-based method to implement the mean-field QM/MM idea of ASEP/MD to plane-wave DFT, which is termed "DFT in classical explicit solvents", or DFT-CES. By employing a three-dimensional real-space grid as a communication medium, we can treat the electrostatic interactions between the DFT solute and the ASEP sampled from MD simulations in a seamless and straightforward manner. Moreover, we couple a fast and efficient free energy calculation method based on the two-phase thermodynamic (2PT) model with our DFT-CES method, which enables direct and simultaneous computation of the solvation free energies as well as the geometric and electronic responses of a solute of interest under the solvation effect. With the aid of DFT-CES/2PT, we investigate the solvation free energies and detailed solvation thermodynamics for 17 types of organic molecules, which show good agreement with the experimental data. We further compare our simulation results with previous theoretical models and assumptions made for the development of implicit solvation models. We anticipate that our proposed method, DFT-CES/2PT, will enable vast utilization of the ASEP/MD method for investigating solvation properties of materials by using periodic DFT calculations in the future.



1. INTRODUCTION

First-principles-based electronic structure theory, such as density functional theory (DFT) calculations, has been employed in a variety of complex systems to elucidate chemical reaction mechanisms,[1,2] chemophysical properties of materials,[3−5] biological processes,[6,7] and many other research fields. Notwithstanding its wide applications and successes, the calculation often assumes the system being in vacuo and excludes the temperature effect by disregarding the dynamics (known as the zero temperature limit) due to its relatively high computational cost. To extend the applicability of first-principles calculations, it is therefore important to develop new methods that can effectively include the dynamic environmental effect without much increasing computational costs.

Solvation effect, which occurs when the system is surrounded by a liquid environment, is of particular interest because the electronic structure and energetics can be substantially modified by complicated solute−solvent interactions. This has led a number of theoretical studies to develop simulation frameworks to include solvation effect in first-principles calculations and to quantify the solvation free energy, which is defined as the thermodynamic changes associated with the solvation process.[8]

Implicit solvation methods have been a popular choice so far and simplify the surrounding solvent by treating it as a homogeneous and continuous dielectric medium.[9,10] To drastically reduce the computational cost, however, it completely neglects the atomistic details of the microscopic solvation structure. Moreover, implicit solvation method conceives the electrostatic response of the media as the leading-order term of the solvation free energy, while the other remaining parts are often parametrized to fit the experimental data of the solvation free energies. This has raised some concerns about the validity of implicit solvent models for systems where nonelectrostatic effects (i.e., hydrophobic effect) are dominant, such as biomolecular systems.[11−13]

On the other hand, explicit solvents include the full atomistic details of the solvent molecules and thereby can properly handle the solvation effect derived from the solvent structure in principle.[14,15] However, it is necessary to sample a high number

Received: May 6, 2016
Published: September 1, 2016

DOI: 10.1021/acs.jctc.6b00469
J. Chem. Theory Comput. 2016, 12, 5088−5099

of solvent degrees of freedom with low symmetry. This often makes the full quantum mechanical (QM) treatment of solute−solvent interactions a computationally impractical choice, and even a hybrid-type solvation method that treats only the solute using QM while using classical approximation for the solvent (so-called QM/MM methods),[16−18] becomes computationally demanding if one attempts to achieve a fully equilibrated state using QM/MM dynamics while considering full solute dynamics.

Meanwhile, the average solvent electrostatic potential/molecular dynamics (ASEP/MD) method developed by the Aguilar group[19,20] provides an efficient and practical route to incorporate the explicit solvent effect into electronic structure calculations by virtue of mean-field type approximations (hence the occasional name "a mean-field QM/MM"[21]). This method is based on the effective Hamiltonian method[22] which provides the theoretical foundation to other common solvation methods such as PCM[23,24] or RISM-SCF,[25] but the solvent field is approximated as an average solvent electrostatic potential (ASEP) sampled from classical MD simulations of explicit solvent molecules. This method has been applied to study electronic excitations,[26,27] transition states,[28,29] solvation free energies,[30,31] and so on. The most advantageous characteristic of this scheme is that it requires only a few QM calculations by neglecting QM dynamics of the solute, yet a large number of solvent configurations can be efficiently treated classically. Additionally, the solute−solvent mutual interaction can be properly described; both the electronic polarization of the solute and the full rotational relaxation of the solvent molecules can be adequately described. However, the technical complication of the ASEP/MD approach lies in designing a communication method between the continuous charge density of the QM solute and the average solvent potential sampled from MD trajectories.

For the realization of ASEP/MD in conjunction with nonperiodic quantum chemistry programs, the simplest method is employing point charges based on the electrostatic potential (ESP) to represent the QM solute in standard MD codes. Also, the ASEP is approximately represented by using a set of fixed point charges.[20] However, this approach occasionally has accuracy and convergence problems depending on the type of systems due to the loss of subtle detailed electrostatic interaction information at short- and long-range. One recently proposed method to adequately treat nonperiodic continuous QM charge density in periodic MD system is based on a rather complex combination of a grid-based treatment for the short-range electrostatic interaction and a point-charge approximation for the long-range electrostatic interaction coupled with a QM/MM-Ewald method.[21]

To expand the applicability of the ASEP/MD method for studying a wide range of solvation effects from molecular to heterogeneous catalytic systems, we herein demonstrate a robust, accurate, and cost-effective implementation to realize the mean-field QM/MM idea coupled with plane-wave DFT (for brevity, our method will be called "DFT in classical explicit solvents", or DFT-CES, hereafter). In DFT-CES, we adopt a real-space rectangular grid which fully occupies a periodic simulation box as a communication medium to seamlessly couple all electrostatic interactions between the QM solute charge density and the MM solvent environment. Moreover, we emphasize that we incorporate the two-phase thermodynamic (2PT) model within the framework of the ASEP/MD approach to calculate solvation free energy by enabling a direct and accurate calculation of the absolute entropy and free-energy quantities of liquids from the MD simulation trajectories.[32]

The outline of this article is as follows. In Section 2, we review and reformulate key equations of ASEP/MD and suggest an algorithm to implement ASEP/MD using a real-space grid. We then discuss the solvation free energy calculation method using the 2PT model. In Section 3, we describe the computational details. In Section 4, we investigate the solvation thermodynamics of 17 organic molecules using DFT-CES calculations. Finally, the paper ends with a short conclusion.

2. BACKGROUND AND THEORY

2.1. Review and Reformulation of ASEP/MD. We consider the solvation effect (sometimes also known as the environment effect mainly described by electrostatic interactions) on the solute by an external electrostatic potential of the solvent, $V_{solvent}$. Under the DFT formalism, we define the effective Hamiltonian of the solute molecule (H_{solute}) as the sum of the original Kohn−Sham (KS) Hamiltonian (H^{KS}) of the solute with no solvation effect and the external electrostatic potential $V_{solvent}$.

$$H_{solute} = H^{KS} + V_{solvent} \tag{1}$$

It can be used to obtain the electron density of the solute, ρ_{solute}, by variationally minimizing the total energy of the solute, $E_{solute}[\rho_{solute}]$ for the given solvent effect of $V_{solvent}$. In general implicit solvation schemes, one approximately evaluates $V_{solvent}$ by considering electrostatic responses of the solvent media for the given solute charge density determined by ρ_{solute}. Due to this mutual dependence between $V_{solvent}$ and ρ_{solute}, both of these can be calculated in an iterative manner, i.e., a self-consistent reaction field (SCRF) method.[33,34] In the ASEP/MD approach, however, $V_{solvent}$ is sampled from classical MD simulation trajectories instead of solving a continuum equation.

The classical Hamiltonian for solvent molecules ($H_{solvent}$) interacting with a solute molecule consists of the classical kinetic energy of the solvent molecules, the interaction energy between solvent molecules (U_{vv}), and the interaction energy between solute and solvent molecules (U_{uv}):

$$H_{solvent} = \sum_i \frac{\mathbf{p}_i^2}{2\mathbf{m}_i} + U_{vv} + U_{uv} \tag{2}$$

where the mass and momentum vector of the ith solvent atom are denoted as $\mathbf{m}_i$ and $\mathbf{p}_i$, respectively.

We use classical force fields (FFs) to describe the interaction energy between solvent molecules (U_{vv}). It should be noted that there exist common classical FFs with relatively high qualities which have been delicately optimized over the decades to reproduce the thermodynamic properties of various solvents.[35−38]

Solute and solvent molecules interact with each other via nonbond interactions (U_{uv}), and thereby we decompose this term into electrostatic interactions (U_{ES}) and van der Waals (vdW) interactions (U_{vdW}) as widely done in classical FF descriptions:

$$U_{uv} = U_{ES} + U_{vdW} \tag{3}$$

We describe the vdW interaction using an atom-position specific potential (mostly 12-6 Lennard−Jones potential) based on the certain choice of FF parameters. Considering that even the current DFT method usually fails to describe the vdW

DOI: 10.1021/acs.jctc.6b00469
J. Chem. Theory Comput. 2016, 12, 5088−5099

interaction and a pairwise additive potential is often used to correct it, our description of U_{vdW} is reasonably acceptable.

U_{ES} describes the electrostatic interaction between the charges of solvent molecules (described using point partial charges of atoms, q_i, in the classical description) and the electron density of the solute (ρ_{solute}) + nuclei charges. If we denote the overall electrostatic potential of the solute as V_{solute}, which consists of Hartree potential and the external nuclei potential, U_{ES} can be written as

$$U_{ES} = \int d^3\mathbf{r}\rho_{solvent}(\mathbf{r})V_{solute}(\mathbf{r}) \tag{4}$$

where the spatial charge distribution $\rho_{solvent}$ induced by the distributed point charges of solvent particles is defined using the Dirac-δ function:

$$\rho_{solvent}(\mathbf{r}) = \sum_i \mathbf{q}_i\delta(\mathbf{r} - \mathbf{r}_i) \tag{5}$$

where $\mathbf{q}_i$ and $\mathbf{r}_i$ denote the partial charge and position vector of the ith atom of the solvent system.

Having so far defined all terms of the classical Hamiltonian of eq 2, we now derive the equations of motion for the solvent, where the force vector exerted on the ith solvent atom $\mathbf{f}_i$ ($\equiv \mathbf{f}_i^0 + \mathbf{f}_i^{ES}$) consists of the force components that can be calculated from the chosen FF parameter $\mathbf{f}_i^0$ and the electrostatic force $\mathbf{f}_i^{ES}$:

$$\mathbf{f}_i^0 = -\nabla_i U_{vv} - \nabla_i U_{vdW} \tag{6}$$

$$\mathbf{f}_i^{ES} = q_i\mathbf{F}_{solute}|_{\mathbf{r}=\mathbf{r}_i} \tag{7}$$

where the electrostatic field generated by the solute molecule, $\mathbf{F}_{solute}$ is defined by

$$\mathbf{F}_{solute}(\mathbf{r}) = -\nabla V_{solute}(\mathbf{r}) \tag{8}$$

For a given DFT result of solute molecule, we can first compute V_{solute} and then perform a classical MD simulation of solvent molecules. From MD simulation trajectories of $\mathbf{r}(t)$ for the time t, we compute the ensemble average of reaction field, $\rho_{solvent}(\mathbf{r})$ that can be replaced with the time average if the solvent dynamics is ergodic:

$$\langle\rho_{solvent}\rangle = \lim_{\tau\to\infty}\frac{1}{\tau}\int_{t=0}^{\tau} dt\rho_{solvent}(\mathbf{r}(t)) \tag{9}$$

By solving the Poisson equation for $\langle\rho_{solvent}\rangle$ sampled from MD simulation trajectories, we finally compute $V_{solvent}$ as the ASEP:

$$\nabla^2 V_{solvent} = -4\pi\langle\rho_{solvent}\rangle \tag{10}$$

which can be plugged into the QM Hamiltonian of eq 1 to obtain an updated ρ_{solute} under the solvation effect. We then perform consecutive iterations of MD and DFT simulations to obtain the updated $V_{solvent}$ (that is plugged into the QM Hamiltonian of the following DFT step) and ρ_{solute} (that is plugged into the MM Hamiltonian of the following MD step), respectively. This allows the consideration of mutual polarization between solute and solvent under the mean-field approximation.Our test for small organic solute molecules has demonstrated that several iteration steps are needed to obtain converged both $V_{solvent}$ and ρ_{solute} (vide infra), which may be further accelerated by employing the multiple environment single system (MESS) QM/MM method in the future.[39,40]

2.2. Implementation: Real-Space Grid-Based Method. In the current method, we exploit a real-space rectangular grid as a seamless communication medium between two different scales of simulations, DFT and MD. Electrostatic interactions between the continuous electron density from DFT and the classical point charges from MD are treated through a background three-dimensional real-space grid. This grid serves as a mapping layer of the electrostatic potential of V_{solute} and $V_{solvent}$.

From a DFT calculation, we compute $\rho_{solute}(\mathbf{r})$ that is converted into V_{solute} combined with the external potential exerted by the nuclei (or pseudopotentials). Using plane-wave DFT codes, V_{solute} is naturally computed on the rectangular real-space grid at a given grid point of $\mathbf{g} = (\Delta_x n_x, \Delta_y n_y, \Delta_z n_z)$ for $n_\alpha = 1, 2, \cdots, N_\alpha$, where Δ_α is the grid spacing, $\alpha \in \{x, y, z\}$, and $N_{grid} = N_x N_y N_z$ is the total number of grid points. The solute geometry and V_{solute} obtained from DFT calculation are placed in the MD simulation box, which are maintained throughout the subsequent MD simulation period (i.e., no dynamics of the solute molecule is assumed). To compute the electrostatic field $\mathbf{F}_{solute}(\mathbf{r})$ used in eq 8, we evaluate the gradient of V_{solute} with the finite-difference method (FDM). Here, we used a first-order FDM, which can be replaced by a higher order method later on.

$$\mathbf{F}_{solute,\alpha}(\mathbf{r})|_{\mathbf{r}=\mathbf{g}} = -\frac{V_{solute}(\mathbf{r})|_{\mathbf{r}=\mathbf{g}+\mathbf{d}_\alpha} - V_{solute}(\mathbf{r})|_{\mathbf{r}=\mathbf{g}-\mathbf{d}_\alpha}}{2|\mathbf{d}_\alpha|} \tag{11}$$

where $\mathbf{d}_x = (\Delta_x, 0, 0)$, $\mathbf{d}_y = (0, \Delta_y, 0)$, and $\mathbf{d}_z = (0, 0, \Delta_z)$. From the DFT simulation side, we therefore obtain four different scalar grid data of $V_{solute}(\mathbf{r})|_{\mathbf{r}=\mathbf{g}}$ and $\mathbf{F}_{solute,\alpha}(\mathbf{r})|_{\mathbf{r}=\mathbf{g}}$ for $\alpha \in \{x, y, z\}$.

To perform MD simulations based on DFT results, we have to compute the potential energy terms and their derivatives, i.e., forces defined in eqs 2, 3, 6, and 7. U_{vv}, U_{vdW}, and $\mathbf{f}_i^0$ can be rather straightforwardly evaluated using a conventional MD simulation code by placing fixed solute atoms at the coordinates of the DFT optimized structure in the previous step (refer Figure 1). To evaluate U_{ES} and $\mathbf{f}_i^{ES}$ using eqs 4 and 7, we need to respectively evaluate $V_{solute}(\mathbf{r})$ and $\mathbf{F}_{solute,\alpha}(\mathbf{r})$ at the off-grid point of $\mathbf{r} = \mathbf{r}_i$, where classical MM particle exists. We thus interpolate V_{solute} and $\mathbf{F}_{solute,\alpha}$ from eight grid points adjacent to arbitrary off-grid atomic position of $\mathbf{r}_i$ using a trilinear interpolation scheme (the full mathematical expressions of such a scheme are given in Figure S1). Because grid data sets (V_{solute} and $\mathbf{F}_{solute,\alpha}$) are prepared only once at the beginning of an MD simulation, these additional electrostatic energy and force evaluation routines follow $O(N_{solvent})$ scale (where $N_{solvent}$ is the total number of solvent atoms in MD simulation), which adds marginal computational cost to the MD simulation part.

During MD simulations, we evaluate $\rho_{solvent}(\mathbf{r})$ on another real-space rectangular grid, followed by its time averaging procedure to yield $\langle\rho_{solvent}\rangle$. For this, we employ a hat function, $\wedge(x)$, as a numerical approximation of the Dirac-δ function of eq 5, smoothing the distributed point partial charges (located at off-grid points) into a continuous charge distribution on grid points;

$$\rho_{solvent}(\mathbf{r}) \cong \sum_i \frac{q_i}{\Delta_x\Delta_y\Delta_z}\wedge\left(\frac{x - x_i}{\Delta_x}\right)\wedge\left(\frac{y - y_i}{\Delta_y}\right)\wedge\left(\frac{z - z_i}{\Delta_z}\right) \tag{12}$$

$$\wedge(x) \equiv \begin{cases} 1 - |x|, & |x| < 1 \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

DOI: 10.1021/acs.jctc.6b00469
J. Chem. Theory Comput. 2016, 12, 5088−5099

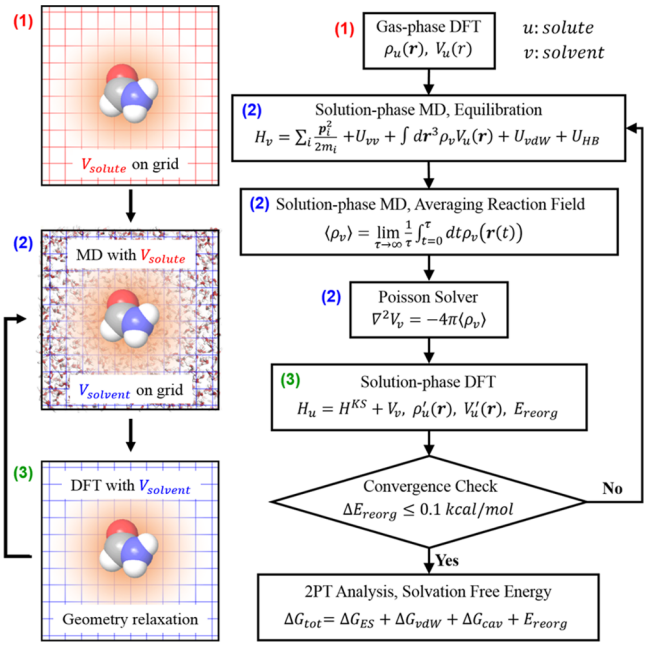


Figure 1. Schematic diagram and flowchart of our DFT-CES method. After a gas-phase DFT calculation, MD simulations are performed with the DFT-generated solute potential, followed by obtaining the ensemble-averaged reaction field from sufficiently long MD trajectories. With the MD-created solvent potential, the geometry and electronic structures of the target solute are updated by a solution-phase DFT calculation. The DFT calculations and MD simulations are iterated until the convergence criterion is met. In the end, further analyses are carried out to extract thermodynamic properties of the system such as solvation free energies.

Because the time complexity of such scheme is also of order $O(N_{solvent})$, the computational burden for both evaluating electrostatic energy and force at the off-grid points and smoothing the point charge located at the off-grid points into the grid-points are independent of the grid size (N_{grid}). This enables us to use a fine enough grid not to lose detailed information for the communication between DFT and MD simulations. After we obtain $\langle \rho_{solvent} \rangle$ from the MD simulation trajectory, we solve the Poisson equation of eq 10 to obtain $V_{solvent}$ on the same dimension of grid. Among a number of algorithms for solving the Poisson equation, a reasonable choice is the one based on fast Fourier transform (FFT) because our simulation cell is periodic, and thus, short- and long-range electrostatic interactions are adequately included without involving Ewald-like algorithms.

For a given $V_{solvent}$, we repeat the DFT calculation using the consecutively updated effective Hamiltonians of eq 1 until convergences of electronic and geometric structures are achieved. Using the plane-wave codes, $\rho_{solvent}$ saved on the grid can be seamlessly added either to the Hartree potential or to the external nuclei potential if the grid dimension matches. Thus, the most straightforward choice would be to employ the same grid dimension to the FFT grid used in the plane-wave DFT code, which is normally in range around 0.1 Å. From the DFT calculation with geometry optimization, we obtain the updated solute geometry, $\rho_{solute}(\mathbf{r})$ and V_{solute}, which can be utilized in the following MD simulation. We then perform consecutive DFT and MD simulations until the electronic and geometric structures of the solute are converged under the solvation effect, i.e., by performing SCRF iterations. The overall computational procedure is summarized as flowchart in Figure 1.

2.3. Solvation Free Energy Calculation. In this section, we discuss how to obtain the total solvation free energy (ΔG_{tot}) from the DFT-CES result converged by full SCRF iterations. Because the solute is treated as a fixed external potential for MD simulation side, we consider the Gibbs free energy change of solvent molecules due to the presence of an external potential of the solute, $\Delta G_{solvent}$. To compute $\Delta G_{solvent}$ directly from the MD simulation trajectory, we use the 2PT model,[32,41,42] which properly approximates the total degree of freedom of the liquid system into the gas-like part (assumed as hard sphere model) and the solid-like part (assumed as harmonic oscillator model), sharing the same viewpoint with Eyring's significant theory of liquids.[43,44] It is noted that the 2PT method has been successful in quantitatively predicting the absolute entropy of various solvents including water,[41,42] solvation free energies from conventional MD simulations,[45,46] surface tensions that are closely related to the cavitation process,[47] and thermodynamics in various other applications.[48,49] Using the 2PT method, we calculated the Gibbs free energies for (a) a bulk solvent system (pure water in this case, $G^0_{solvent}$), and (b) a solution system with the solute external potential obtained after full SCRF iterations ($G_{solvent}$). The difference between (a) and (b) gives us quantitative information about thermodynamic changes of the solvent system during the solvation process ($\Delta G_{solvent}$) as:

$$\Delta G_{solvent} = G_{solvent} - G^0_{solvent} \tag{14}$$

By extension, to analyze elemental components of $\Delta G_{solvent}$ (electrostatic, dispersion, and cavitation free energy terms), based on the final converged solute electronic and geometric structures, we performed two additional (nonself-consistent) DFT-CES/2PT simulations by sequentially switching off electrostatic and vdW interactions between solute and solvent. $G_{solvent}|_{noES}$ is defined as the solvent free energy calculated when only the solute−solvent electrostatic interaction is turned off. $G_{solvent}|_{WCA}$ is defined as the solvent free energy calculated when only the solute−solvent repulsive interaction is left (without including electrostatic and dispersive interactions, which is accomplished by employing the Weeks−Chandler−Anderson (WCA) repulsive potential.[50] These lead to the definition of electrostatic, dispersion, and cavitation free energy terms as

$$\Delta G_{ES} \equiv G_{solvent} - G_{solvent}|_{noES} \tag{15}$$

$$\Delta G_{disp} \equiv G_{solvent}|_{noES} - G_{solvent}|_{WCA} \tag{16}$$

$$\Delta G_{cav} \equiv G_{solvent}|_{WCA} - G^0_{solvent} \tag{17}$$

respectively, and then yielding

$$\Delta G_{solvent} = \Delta G_{ES} + \Delta G_{disp} + \Delta G_{cav} \tag{18}$$

On the other hand, because the DFT Hamiltonian of eq 1 treats the solvent effect as an external potential, the total DFT energy of the solute (E_{solute}) consists of internal energy of the solute (E_{int}) and solute−solvent electrostatic interaction energy (E_{ES}). As in the previous methods,[51,52] we then define the reorganization energy of the solute (E_{reorg}) as the internal energy change associated with the electronic and geometric structure changes of the solute molecule in response to the solvent:

$$E_{reorg} = E_{int} - E^0_{solute} \tag{19}$$

DOI: 10.1021/acs.jctc.6b00469
J. Chem. Theory Comput. 2016, 12, 5088−5099

where E^0_{solute} is calculated as the energy of the solute fully optimized in vacuo. Because the solute−solvent electrostatic interaction energy (E_{ES}) is already treated in the MD side (ΔG_{ES}), we only have taken E_{reorg} into the calculation of the total solvation free energy as

$$\Delta G_{tot} = \Delta G_{solvent} + E_{reorg} \quad (20)$$

where the first and second terms can be obtained from MD and DFT simulations, respectively.

3. COMPUTATIONAL DETAILS

3.1. Method Implementation. We implemented the DFT-CES method by using open-source DFT and MD programs: the plane-wave DFT code of Quantum ESPRESSO,[53] and the highly modular and massively parallelized MD code of LAMMPS: Large-scale Atomic/Molecular Massively Parallel Simulator.[54] We modified the built-in subroutine of Quantum ESPRESSO to read the grid data of $V_{solvent}$ and add them to the calculated Hartree potential during self-consistent field cycles. We also modified energy and force evaluation subroutines for geometry optimization under an external potential. LAMMPS provides full functional library which can be linked with external driver programmed by C/C++, FORTRAN, and Python languages. Thus, we built an external C++ program linked with LAMMPS library which interpolates the V_{solute} and $\mathbf{F}_{solute,\alpha}$ grid data on-the-fly for any arbitrary position of MD particles and then computes the electrostatic energy and forces during the regular time evolution of MD simulation. This external program simultaneously computes $\rho_{solvent}$ on the grid using eq 12 and collects it during MD simulation to compute its time average for the evaluation of ASEP, $\langle\rho_{solvent}\rangle$.

3.2. Simulation Setup and Systems. Using the DFT-CES/2PT method, we calculated hydration free energies of 17 types of organic molecules listed in Figure S2. It is noted that previous studies have tested this set of molecules using various solvation models, including classical MD with alchemical free energy calculations (explicit solvent),[55] Poisson−Boltzmann (PB) with model charges (implicit solvent),[55] and DFT with the SM8 implicit solvation model.[56]

For DFT calculations, we chose the Perdew−Burke−Ernzerhof (PBE) exchange-correlation functional,[57] and electron−ion interactions were considered in the form of the projector-augmented-wave (PAW) method with a plane wave up to an energy cutoff of 650 eV. For MD simulations, we performed canonical ensemble (NVT) simulations at 300 K with the Nosé−Hoover thermostat.[58,59] The simulation cell contains 1000 TIP3P[60] water molecules as well as a fixed single solute molecule whose geometry was obtained from the previous DFT optimization step. During the simulation, all of the water molecules were kept rigid by means of the SHAKE[61] algorithm. The vdW parameters for solute molecules were taken from the OPLS-AA FF[35] parameter set. A global cutoff distance of 15 Å for nonbonded interaction was used with the Ewald summation method[62] for long-range electrostatic interactions. Aside from the water-solute system, in advance, we performed the MD simulation of a bulk water system containing 1000 water molecules to compute the free energy of $G^0_{solvent}$ which serves as the reference.

Combining the DFT and MD procedures at each iteration step for the SCRF method, we obtained the fully optimized geometry and electronic structure of a target solute on the DFT side and then carried out 3 ns MD simulations with a time step of 1 fs. We averaged the solvent charge density at every time step on-the-fly for the last 1 ns dynamics to evaluate $\langle\rho_{solvent}\rangle$ on the MD side. It is noted that the simulation setup consisting of 2 ns equilibration followed by 1 ns ASEP calculation is rather a conservative choice to fully equilibrate the system (cf. the rotational relaxation time scale of the bulk water is in picosecond order[63]), which may be further optimized to reduce the entire computational cost.

3.3. vdW Interaction of Atoms Involved in Hydrogen Bond Interaction. The OPLS-AA and the modified TIP3P water FFs describe the hydrogen bond (HB), which is relatively stronger than other conventional intermolecular interactions, by switching off vdW repulsion of an HB donating hydrogen atom, i.e., $U_{vdW} = 0$. Such a seemingly unphysical setting for vdW parameters accompanied by a well-balanced partial point-charge description is empirically adequate, thereby properly reproducing bulk properties of various liquid systems. However, when the realistic DFT electron density information is incorporated into the solute system, the short-range electrostatic interaction will be changed due to the smeared electron density compared with the classical point-charge description. This leads to systematic overestimation of solvation free energies (Table S1). It is thus required to set a finite U_{vdW} for the HB interacting atoms to achieve a more accurate description.

To describe U_{vdW} for the HB interacting atoms, we used the explicit 3-body Morse potential term in eq 21, which was originally designed for describing HB interaction in DREIDING FFs by Goddard and co-workers.[64] By employing such a relatively short-range potential having no algebraic tail, one can recover the original OPLS-AA description of dispersive interaction at the large separation limit, where the electron density smearing effect is neglected. The use of Morse potential is further advantageous in terms of a mathematical flexibility because it has three adjustable parameters:

$$U_{HB} = \begin{cases} D_0\{e^{-2\gamma(r-r_0)} - 2e^{-\gamma(r-r_0)}\}\cos^2\theta, & |\theta| > \pi/2 \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

where the acceptor−donor distance r and the acceptor-hydrogen-donor bond angle θ defined in the range of $[-\pi, \pi]$, as illustrated in Figure S3-a. Three parameters, D_0, r_0, and γ in eq 21 were determined by fitting to the high level CCSD(T) binding energy data sets for hydrogen bonding pairs[65] with single-point DFT-CES calculation. More details and the fitted parameters are provided in Figures S3b−l and Table S2 in the Supporting Information.

4. RESULTS AND DISCUSSION

4.1. Convergence of SCRF Iterations. To check the convergence of DFT-CES, we monitored how E_{reorg} from DFT calculations changed over the SCRF iterations. As shown in Figure 2a, the difference of E_{reorg} between two consecutive iterations was converged to less than 0.1 kcal/mol within less than 10 SCRF iterations for all types of 17 solute cases. This shows that the overall DFT-CES computational time is limited by the several sets of DFT geometry optimizations of the solute molecule as well as several sets of NVT MD simulations of the solvent box.

In principle, the mutual solute−solvent electrostatic interaction energy should be identical no matter whether the energy

DOI: 10.1021/acs.jctc.6b00469
J. Chem. Theory Comput. 2016, 12, 5088−5099

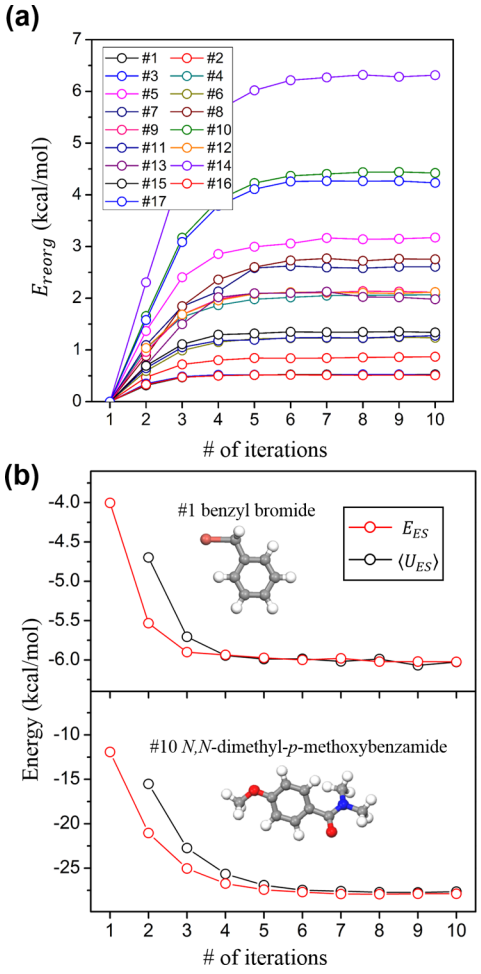


Figure 2. Convergence trend of (a) the solute reorganization energies (E_{reorg}) from DFT and (b) the solute−solvent electrostatic interaction energies evaluated from DFT (E_{ES}) and MD ($\langle U_{ES} \rangle$) in terms of DFT-CES iteration steps. The electronic structures of all 17 solutes and the mutual polarization effect between solute and solvent systems are stably converged within 10 iteration steps.

is calculated from the DFT or the MD side. We thus also monitored the change in E_{ES} from DFT calculations as well as the ensemble-averaged $\langle U_{ES} \rangle$ from MD simulations (defined in eq 4) over the SCRF iterations (representative cases are shown in Figure 2b), which shows good convergence. We note that E_{ES} and $\langle U_{ES} \rangle$ agree with each other within the mean absolute error (MAE) and root-mean-square error (RMSE) values of 0.27 and 0.33 kcal/mol, respectively, as shown in Table S3.

4.2. Reaction Field Sampled from MD. After running enough SCRF iterations to converge E_{reorg}, we visualized the reaction field of $\langle \rho_{solvent} \rangle$ in Figure 3 for two representative cases of ethylene glycol diacetate (#12) and imidazole (#14). The reaction field is clearly shown to have been adjusted to the polarized electron density from DFT calculations. It is particularly noteworthy that the reaction field shows realistic multiple ± fluctuations originating from the local water structure near the solute molecule, in that capturing such features becomes available by considering only explicit solvents (comparison with implicit PB reaction field in Figure S4). Moreover, it is clearly shown that point partial charges of solvent molecules are smoothly smeared out during the dynamics of solvent particles, which is advantageous in avoiding a polarization catastrophe.

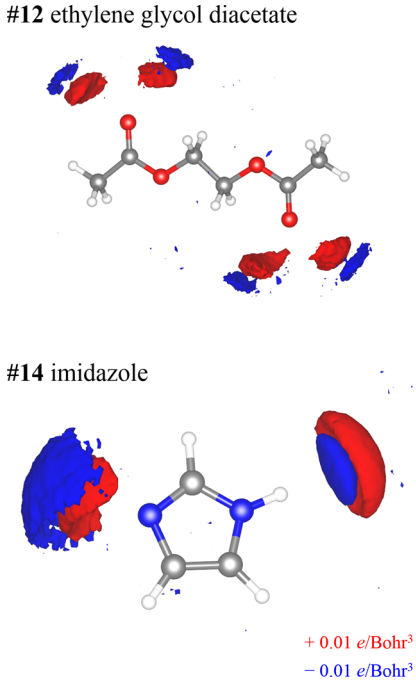


Figure 3. Isosurface of the solvent charge density from SCRF of #12 and #14 solute molecules. It clearly shows that the polarized reaction fields are adjusted to the accurate DFT electron density.

4.3. Solvation Free Energies. With the aid of DFT-CES/2PT calculations, the solvation free energies (ΔG_{tot}) of 17 solutes are listed in Table 1 in comparison with the experimental data.[55] The MAE and RMSE values are given as 1.32 and 1.57 kcal/mol, respectively. We find that the calculated ΔG_{tot} of two bulky benzamide solutes, *N,N*-dimethyl-*p*-methoxybenzamide (#10) and *N,N,*4-trimethylbenzamide (#17), are less negative (i.e., less favoring their

Table 1. List of Hydration Free Energies for 17 Organic Solutes Derived from the Experiments and the DFT-CES/2PT Method[a]

solute molecule	experimental	DFT-CES
#1 benzyl bromide	−2.38	−2.32 (0.45)
#2 benzyl chloride	−1.93	−3.25 (0.49)
#3 bis(2-chloroethyl) ether	−4.23	−2.49 (0.43)
#4 1,1-diacetoxyethane	−4.97	−5.98 (0.43)
#5 1,1-diethoxyethane	−3.28	−4.12 (0.46)
#6 1,2-diethoxyethane	−3.54	−1.57 (0.45)
#7 diethyl propanedioate	−6.00	−6.18 (0.43)
#8 diethyl sulfide	−1.43	−3.95 (0.47)
#9 dimethoxymethane	−2.93	−3.27 (0.47)
#10 *N,N*-dimethyl-*p*-methoxybenzamide	−11.01	−8.70 (0.45)
#11 1,4-dioxane	−5.05	−3.15 (0.44)
#12 ethylene glycol diacetate	−6.34	−6.17 (0.46)
#13 glycerol triacetate	−8.84	−8.58 (0.46)
#14 imidazole	−9.81	−11.46 (0.41)
#15 phenyl formate	−3.82	−5.61 (0.45)
#16 *m*-bis(trifluoromethyl) benzene	1.07	−0.76 (0.43)
#17 *N,N,*4-trimethylbenzamide	−9.76	−7.15 (0.42)
MAE		1.32
RMSE		1.57

[a]The standard deviation values are given in parentheses. The MAE and RMSE values are also included in kcal/mol.

DOI: 10.1021/acs.jctc.6b00469
J. Chem. Theory Comput. 2016, 12, 5088−5099

solvation) than the experimental values by a relatively large amount (2.31 and 2.61 kcal/mol, respectively). On the other hand, the calculated ΔG_{tot} of diethyl sulfide (#8) is more negative (i.e., more favoring its solvation) than the experimental values by 2.52 kcal/mol. If these three solutes are excluded, the MAE and RMSE values are reduced to 1.07 and 1.29 kcal/mol, respectively.

We note that the accuracy of the DFT-CES/2PT method is comparable to that of the alchemical free energy calculations using classical MD simulations reported from the previous study, which exhibits the RMSE spanning from 1.33 to 2.05 kcal/mol depending on the charge schemes.[55] Moreover, the calculated ΔG_{tot} shows an accuracy slightly higher than that of the PB implicit solvation method with nonpolarizable point charges with a single conformer, having an RMSE of 1.87−2.57 kcal/mol depending on the choice of vdW radii.[55]

Compared to that of sophisticatedly developed implicit solvation methods coupled with DFT (polarizable charge density), the current explicit method exhibits a rather lower accuracy, as generally shown in other solvation free energy benchmarking data;[66−68] the MAE and RMSE of the SM8 model are 0.88 and 1.08 kcal/mol, respectively.[56] We note that the outliers of these calculations are still the two bulky benzamide solutes (#10 and #17). As in our DFT-CES method, those imply the possible existence of a systematic error included in the calculation of ΔG_{tot} for those systems. We further want to mention that the implicit solvent models were developed by fitting atomic radii and solute−solvent interfacial free energies to experimental solvation free energies. Therefore, the accuracy level of the current method (which is already acceptable with no further parameter fitting procedure targeted for the experimental solvation free energies) could be further enhanced by delicately tuning the vdW parameters describing solute−solvent interactions.

4.4. Component Breakdown and Enthalpy−Entropy Interplay for Solvation Free Energies. For in-depth analyses of the calculated solvation free energies (ΔG_{tot}), we decomposed ΔG_{tot} for 17 solutes into four types of energies: (a) reorganization energy (E_{reorg}) obtained from DFT calculations, (b) electrostatic, (c) dispersion, and (d) cavitation free energies computed from MD (ΔG_{ES}, ΔG_{disp}, and ΔG_{cav} as defined in eqs 15−17). These values are used to understand the solvation thermodynamics, and the results are shown in Table 2.

Figure 4 shows the linear dependence of the solvation free energy components of ΔG_{ES} and ΔG_{disp} on their ensemble-averaged corresponding interaction energies of $\langle U_{ES}\rangle$ and $\langle U_{vdW}\rangle$, respectively. Using the least-squared fitting, we obtain $\Delta G_{ES} \approx 0.40\langle U_{ES}\rangle$ and $\Delta G_{disp} \approx 0.75\langle U_{vdW}\rangle$. Considering the process of turning on solute−solvent interactions from $\langle U_{uv}\rangle_0$ to $\langle U_{uv}\rangle$, the linear-response theory leads the solvation free energy to be $\Delta G_{solvent} = (\langle U_{uv}\rangle + \langle U_{uv}\rangle_0)/2$. Turning on the electrostatic interaction is expected to gradually reorient the dipoles of water molecules, resulting in $\langle U_{ES}\rangle_0 = 0$ and a finite value of $\langle U_{ES}\rangle$;[69] thus, the linear-response theory gives us the equality of $\Delta G_{ES} = 0.5\langle U_{ES}\rangle$, which is often assumed in implicit solvation models. This is considerably consistent with our numerical results, which further suggest use of the scaling coefficient of 0.4 instead of 0.5, presumably due to the slightly nonlinear response to the electrostatic interaction, although further numerical tests would be required for extensive sets. Turning on the vdW interaction is expected to be insensitive to the orientations of water molecules, resulting in $\langle U_{vdW}\rangle_0 = \langle U_{vdW}\rangle$.[69] This leads to $\Delta G_{disp} = \langle U_{vdW}\rangle$ using linear-response theory in reasonable agreement with our numerical results.

Table 2. Free Energy Components of the Hydration Free Energy from the DFT-CES/2PT Method (kcal/mol)[a]

solute molecule	ΔG_{tot}	E_{reorg}	ΔG_{ES}	ΔG_{disp}	ΔG_{cav}	ΔG_{nonES}
#1	−2.32	0.53	−2.93	−12.49	12.58	0.09
#2	−3.25	0.87	−3.34	−12.52	11.75	−0.77
#3	−2.49	0.52	−3.25	−13.13	13.37	0.24
#4	−5.98	2.07	−7.45	−14.36	13.76	−0.60
#5	−4.12	3.17	−7.92	−14.30	14.92	0.62
#6	−1.57	1.25	−3.91	−10.60	11.69	1.09
#7	−6.18	2.61	−7.01	−16.20	14.42	−1.78
#8	−3.95	2.76	−6.45	−11.55	11.29	−0.25
#9	−3.27	2.12	−6.57	−11.37	12.54	1.17
#10	−8.70	4.44	−10.35	−19.56	16.78	−2.78
#11	−3.15	1.29	−2.73	−11.19	9.49	−1.70
#12	−6.17	2.11	−7.94	−13.89	13.54	−0.34
#13	−8.58	2.91	−12.48	−18.55	19.54	0.98
#14	−11.46	6.31	−17.38	−8.77	8.38	−0.40
#15	−5.61	1.34	−4.87	−14.72	12.64	−2.08
#16	−0.76	0.51	−3.22	−13.75	15.70	1.95
#17	−7.15	4.26	−10.28	−17.20	16.06	−1.14

[a] $\Delta G_{tot} = E_{reorg} + \Delta G_{solvent}$; $\Delta G_{solvent} = \Delta G_{ES} + \Delta G_{disp} + \Delta G_{cav}$; $\Delta G_{nonES} = \Delta G_{disp} + \Delta G_{cav}$.

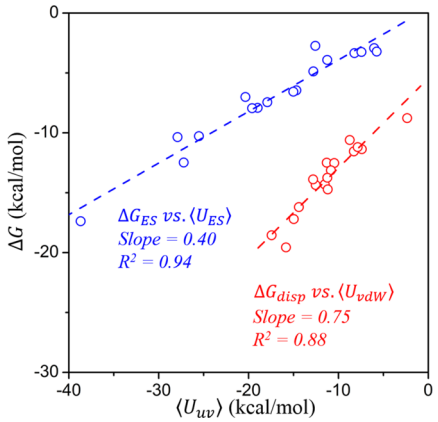


Figure 4. Comparison of the electrostatic component of the solvation free energy, ΔG_{ES}, and the ensemble-averaged solute−solvent electrostatic interaction energy, $\langle U_{ES}\rangle$ (blue), and comparison of the dispersion component of the solvation free energy, ΔG_{disp}, and the ensemble-averaged solute−solvent van der Waal's interaction energy $\langle U_{vdW}\rangle$ (red).

The contribution of nonelectrostatic effect to the solvation free energy ($\Delta G_{nonES} = \Delta G_{disp} + \Delta G_{cav}$; refer to Table 2) is often conceived to be proportional to the solvent accessible surface area (ASA) or the molecular volume of solutes in various implicit solvent models.[33,70−74] However, as shown in Figure 5, there exists almost zero correlation between ΔG_{nonES} and ASA (R^2 = −0.06). On the other hand, the separate components of ΔG_{disp} and ΔG_{cav} scale as a function of ASA with well-defined negative (slope = −28.64 mN/m; R^2 = 0.82) and positive (slope =27.79 mN/m; R^2 = 0.87) linear correlations, respectively. Despite the differences in the technical details during energy decomposition, the same tendency has been reported in previous literatures[75−78] in which both the attractive (leading to a negative slope) and repulsive (leading to a positive slope) components individually

DOI: 10.1021/acs.jctc.6b00469
J. Chem. Theory Comput. 2016, 12, 5088−5099

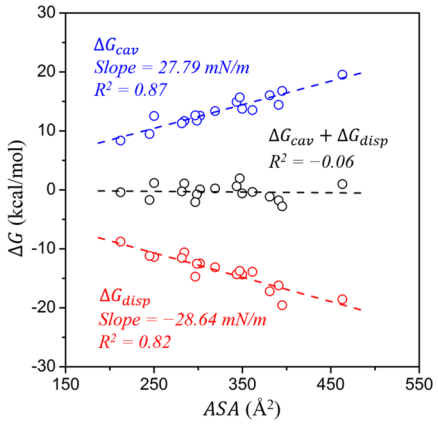


Figure 5. Correlations between nonelectrostatic solute−solvent interaction free energy terms (ΔG_{cav} and ΔG_{disp}) and solvent accessible surface area (ASA).

correlate with ASA, but they are canceled by each other, resulting in a poor correlation when they are added. Therefore, our results also support the existing concerns about a potential limitation of implicit solvation method based on a surface area in treating nonelectrostatic components at once.[13,79−82] We note that a similar tendency have been observed for the correlation between nonelectrostatic component and molecular volume, as shown in Figure S5.

We further discuss the slope of the correlation line between ΔG_{cav} and ASA, which is related to the surface tension of the microscopic cavity, γ. We obtain γ = 27.79 mN/m from least-squared fit, which is substantially smaller than the bulk surface tension of water, γ_{bulk} (52.6 mN/m from MD calculation[47] and 71.99 mN/m from experiment[83]). This is explained using the microscopic cavitation theory formulated by Chandler and his co-workers,[84] predicting $\gamma \approx \gamma_{bulk}(1 - 2\delta/R)$,[85] where δ is the Tolman length and R is the radius of the cavity. Considering $\delta \sim 1$ Å for water and R = 4−6 Å for our solutes, the free energetic cost is expected to decrease by 33−50% for the formation of microscopic cavity compared with the formation of macroscopic surface, as the water molecules near the interface do not need to break but simply reorder the hydrogen bonds, as fully discussed in the previous theoretical works.[84,86]

One of the merits of employing the 2PT method for free energy calculations is the direct computability of the enthalpic and entropic contributions to the free energy quantities. Figure 6a shows the enthalpic (ΔH_{tot}) and entropic ($-T\Delta S_{tot}$) contributions to the solvation free energy (ΔG_{tot}), exhibiting a well-known enthalpy−entropy compensation behavior (R^2 = 0.54 or 0.77 without imidazole).[87] We find one significantly anomalous case for the solvation thermodynamics of imidazole, where the enthalpic contribution dominates the entropic contribution; ΔH_{tot} = −14.75 kcal/mol vs $-T\Delta S_{tot}$ = 3.29 kcal/mol. We conceive that the origin of such an atypical behavior of solvation thermodynamics of imidazole is due to its extraordinarily high polarity (dipole moment = 3.61 D) concentrated within a relatively small molecular volume, resulting in a strong solute−solvent electrostatic interaction.

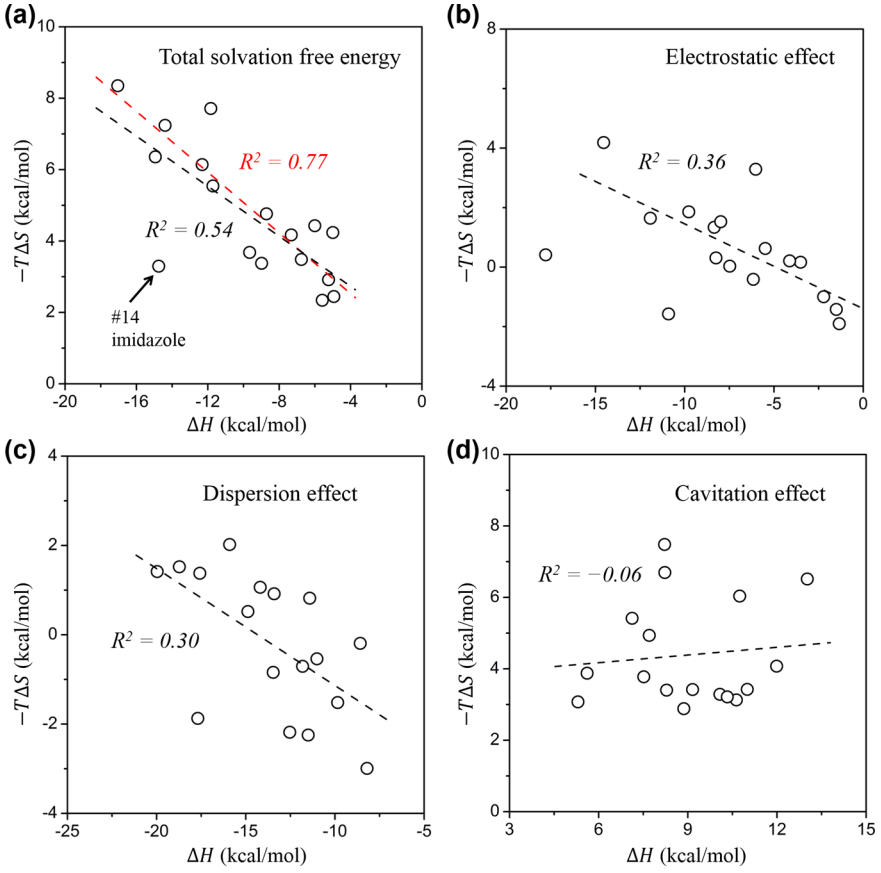


Figure 6. Enthalpy−entropy compensation behaviors for (a) total solvation free energy, (b) electrostatic interaction free energy, (c) dispersion interaction free energy, and (d) cavitation free energy.

DOI: 10.1021/acs.jctc.6b00469
J. Chem. Theory Comput. 2016, 12, 5088−5099

We further investigate the enthalpic (ΔH) and entropic ($-T\Delta S$) contributions to each of ΔG_{ES}, ΔG_{disp}, and ΔG_{cav} (Table S4). For individual terms, we interestingly find only a very weak enthalpy−entropy compensation behavior in the case of electrostatic and dispersion interaction (Figures 6b and c), while the correlation between entropy and enthalpy for cavitation is observed to be almost nonexistent (Figure 6d). Although there exists a certain degree of enthalpy−entropy compensation behavior, it is further noted that the absolute extent of entropy contributions to electrostatic and dispersion interactions are marginal, and thus both interactions are mostly dominated by enthalpic gain.

Figure 7 displays the enthalpic and entropic contributions to nonelectrostatic interactions as a function of ASA, and the

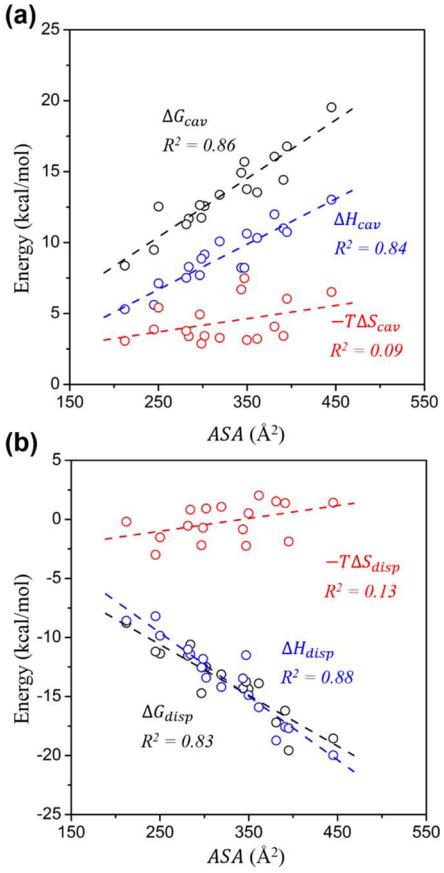


Figure 7. Enthalpic and entropic contributions to the nonelectrostatic solvation free energy components (a) ΔG_{cav} and (b) ΔG_{disp} with respect to the ASA.

entropic contribution to ΔG_{disp} is marginal, while the entropic contribution to ΔG_{cav} is no longer negligible, although the cavitation is still an enthalpy-driven process. This leads to a simplified conclusion that the entire solvation entropy (ΔS_{tot}) is mostly determined by the cavitation entropy (ΔS_{cav}), which has been employed to explain the counterintuitive behavior of solvation thermodynamics: hydration entropies of nonpolar, polar, and charged molecules show strikingly small differences if they have similar sizes.[69] However, it should also be noted that the direct comparison of ΔS_{tot} and ΔS_{cav} shows a relatively weak correlation (Figure 8), implying that the above conclusion might be only qualitatively valid.

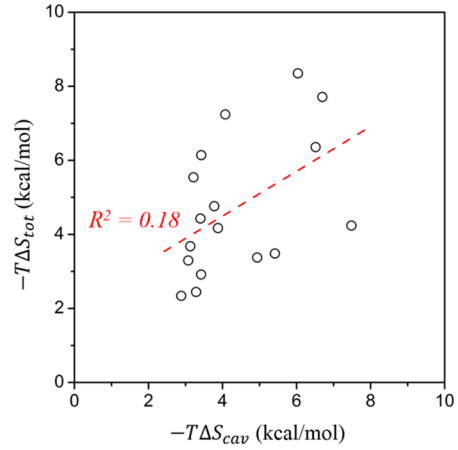


Figure 8. Comparison of entropic contributions between total solvation free energies and cavitation free energies.

5. CONCLUSIONS AND OUTLOOK

In this study, we demonstrate a real-space grid-based method to implement the mean-field coupling concept of the ASEP/MD method, namely DFT-CES. Our basic motivations here are (1) to enable ASEP/MD calculations using periodic DFT code with plane-wave basis sets, and also (2) to couple the fast and accurate free energy calculation method of 2PT with the ASEP/MD formalism for the calculation of solvation thermodynamic properties.

On the real-space rectangular grid with the same dimensions of the grid used in plane-wave DFT calculations, we seamlessly and accurately handle the mutual electrostatic interaction between the DFT charge density and ASEP from MD simulations. We also emphasize the computational efficiency of our algorithm: (1) the additional computational costs from the DFT side are marginal because the ASEP is simply added to the Hartree potential on the grid, and (2) the additional computational complexity to the MD code is $O(N_{solvent})$. Through the adoption of a trilinear interpolation scheme and a smoothing version of the Dirac-δ function, it is also independent of the grid dimension. This allows use of sufficiently fine grids to achieve high accuracy without significantly increasing the computational costs.

Coupling of the 2PT method with DFT-CES further demonstrates the availability of direct calculation of solvation free energies from ASEP/MD results without involving complicated and computationally expensive advanced sampling techniques such as thermodynamic integration (TI), free energy perturbation (FEP), potential of mean force (PMF) calculations, and so on. The benchmark of our method on the solvation of small organic molecules demonstrated good agreement with the experimental values, and individual components of the solvation free energy were decomposed and analyzed in comparison with previous theoretical solvation models.

On the basis of the successful implementation of ASEP/MD, which includes the full atomistic details of the solvent structure in plane-wave DFT, we expect its widespread application in the exploration of materials where the solvent structure changes significantly at the solute−solvent interface. For example, applying our DFT-CES methodology to investigate heterogeneous catalysts could be of great interest because the reliability of implicit solvation models on this heterogeneous interface has

DOI: 10.1021/acs.jctc.6b00469
J. Chem. Theory Comput. 2016, 12, 5088−5099

not been fully addressed yet. Therefore, we anticipate that the increased applicability of periodic DFT with ASEP/MD can provide an efficient and accurate route to study various interesting problems occurring at solid−liquid interfaces.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.6b00469.

Mathematical details of the trilinear interpolation scheme and supplementary figures and tables (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: linus16@kaist.ac.kr.

Author Contributions

§H.-K.L. and H.L. contributed equally to this work.

Funding

This work was supported by the Technology Development Program to Solve Climate Changes (Grant 2015M1A2A2056561) and Nanomaterial Technology Development Program (Grant 2016M3A7B4025414) through the National Research Foundation (NRF) funded by the Ministry of Science, ICT & Future Planning.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge the Supercomputing Center/Korea Institute of Science and Technology Information with supercomputing resources (KSC-2015-C3-005). We also sincerely thank our colleague Prof. Dario Rocca at Université de Lorraine who helped us with the initial implementation of the DFT-CES scheme in the Quantum ESPRESSO program.

■ REFERENCES

(1) Bell, A. T.; Head-Gordon, M. Quantum Mechanical Modeling of Catalytic Processes. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2*, 453−477.

(2) Tsipis, A. C. DFT flavor of coordination chemistry. *Coord. Chem. Rev.* **2014**, *272*, 1−29.

(3) Cramer, C. J.; Truhlar, D. G. Density functional theory for transition metals and transition metal chemistry. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10757−10816.

(4) Neugebauer, J.; Hickel, T. Density functional theory in materials science. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, *3*, 438−448.

(5) Norskov, J. K.; Abild-Pedersen, F.; Studt, F.; Bligaard, T. Density functional theory in surface chemistry and catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 937−943.

(6) LaPointe, S. M.; Weaver, D. F. A review of density functional theory quantum mechanics as applied to pharmaceutically relevant systems. *Curr. Comput.-Aided Drug Des.* **2007**, *3*, 290−296.

(7) Siegbahn, P. E. M.; Blomberg, M. R. A. Density functional theory of biologically relevant metal centers. *Annu. Rev. Phys. Chem.* **1999**, *50*, 221−249.

(8) Skyner, R. E.; McDonagh, J. L.; Groom, C. R.; van Mourik, T.; Mitchell, J. B. O. A review of methods for the calculation of solution free energies and the modelling of systems in solution. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6174−6191.

(9) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378−6396.

(10) Mathew, K.; Sundararaman, R.; Letchworth-Weaver, K.; Arias, T. A.; Hennig, R. G. Implicit solvation model for density-functional study of nanocrystal surfaces and reaction pathways. *J. Chem. Phys.* **2014**, *140*, 084106.

(11) Zhou, R. H.; Berne, B. J. Can a continuum solvent model reproduce the free energy landscape of a beta-hairpin folding in water? *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12777−12782.

(12) Zhou, R. H.; Krilov, G.; Berne, B. J. Comment on "Can a continuum solvent model reproduce the free energy landscape of a beta-hairpin folding in water?" The Poisson-Boltzmann equation. *J. Phys. Chem. B* **2004**, *108*, 7528−7530.

(13) Mennucci, B.; Tomasi, J.; Cammi, R.; Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Gabriel, S.; Stephens, P. J. Polarizable Continuum Model (PCM) Calculations of Solvent Effects on Optical Rotations of Chiral Molecules. *J. Phys. Chem. A* **2002**, *106*, 6102−6113.

(14) Shivakumar, D.; Deng, Y. Q.; Roux, B. Computations of Absolute Solvation Free Energies of Small Molecules Using Explicit and Implicit Solvent Model. *J. Chem. Theory Comput.* **2009**, *5*, 919−930.

(15) Wu, X. W.; Sung, S. S. Simulation of peptide folding with explicit water - A mean solvation method. *Proteins: Struct., Funct., Genet.* **1999**, *34*, 295−302.

(16) Friesner, R. A.; Guallar, V. Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389−427.

(17) Liu, M. Y.; Wang, Y. J.; Chen, Y. K.; Field, M. J.; Gao, J. L. QM/MM through the 1990s: The First Twenty Years of Method Development and Applications. *Isr. J. Chem.* **2014**, *54*, 1250−1263.

(18) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.

(19) Galvan, I. F.; Sanchez, M. L.; Martin, M. E.; del Valle, F. J. O.; Aguilar, M. A. ASEP/MD: A program for the calculation of solvent effects combining QM/MM methods and the mean field approximation. *Comput. Phys. Commun.* **2003**, *155*, 244−259.

(20) Sanchez, M. L.; Aguilar, M. A.; delValle, F. J. O. Study of solvent effects by means of averaged solvent electrostatic potentials obtained from molecular dynamics data. *J. Comput. Chem.* **1997**, *18*, 313−322.

(21) Nakano, H.; Yamamoto, T. Accurate and Efficient Treatment of Continuous Solute Charge Density in the Mean-Field QM/MM Free Energy Calculation. *J. Chem. Theory Comput.* **2013**, *9*, 188−203.

(22) Tomasi, J.; Bonaccorsi, R.; Cammi, R.; Delvalle, F. J. O. Theoretical Chemistry in Solution - Some Results and Perspectives of the Continuum Methods and in Particular of the Polarizable Continuum Model. *J. Mol. Struct.: THEOCHEM* **1991**, *234*, 401−424.

(23) Mennucci, B. Polarizable continuum model. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 386−404.

(24) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999−3093.

(25) Kovalenko, A.; Hirata, F. Three-dimensional density profiles of water in contact with a solute of arbitrary shape: a RISM approach. *Chem. Phys. Lett.* **1998**, *290*, 237−244.

(26) Galvan, I. F.; Martin, M. E.; Munoz-Losa, A.; Sanchez, M. L.; Aguilar, M. A. Solvent Effects on the Structure and Spectroscopy of the Emitting States of 1-Phenylpyrrole. *J. Chem. Theory Comput.* **2011**, *7*, 1850−1857.

(27) Losa, A. M.; Galvan, I. F.; Martin, M. E.; Aguilar, M. A. Solvent effects on the low-lying excited states of a model of retinal. *J. Phys. Chem. B* **2006**, *110*, 18064−18071.

(28) Corchado, J. C.; Sanchez, M. L.; Galvan, I. F.; Martin, M. E.; Munoz-Losa, A.; Barata-Morgado, R.; Aguilar, M. A. Theoretical Study of Solvent Effects on the Ground and Low-Lying Excited Free Energy Surfaces of a Push-Pull Substituted Azobenzene. *J. Phys. Chem. B* **2014**, *118*, 12518−12530.

(29) Galvan, I. F.; Martin, M. E.; Aguilar, M. A. A new method to locate saddle points for reactions in solution by using the free energy gradient method and the mean field approximation. *J. Comput. Chem.* **2004**, *25*, 1227−1233.

(30) Kamerlin, S. C. L.; Haranczyk, M.; Warshel, A. Progress in Ab Initio QM/MM Free Energy Simulations of Electrostatic Energies in

Proteins: Accelerated QM/MM Studies of pK(a), Redox Reactions and Solvation Free Energies. *J. Phys. Chem. B* **2009**, *113*, 1253−1272.
(31) Rosta, E.; Haranczyk, M.; Chu, Z. T.; Warshel, A. Accelerating QM/MM free energy calculations: Representing the surroundings by an updated mean charge distribution. *J. Phys. Chem. B* **2008**, *112*, 5680−5692.
(32) Lin, S. T.; Blanco, M.; Goddard, W. A. The two-phase model for calculating thermodynamic properties of liquids from molecular dynamics: Validation for the phase diagram of Lennard-Jones fluids. *J. Chem. Phys.* **2003**, *119*, 11792−11805.
(33) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. New model for calculation of solvation free energies: Correction of self-consistent reaction field continuum dielectric theory for short-range hydrogen-bonding effects. *J. Phys. Chem.* **1996**, *100*, 11775−11788.
(34) Tapia, O.; Goscinski, O. Self-Consistent Reaction Field-Theory of Solvent Effects. *Mol. Phys.* **1975**, *29*, 1653−1661.
(35) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(36) Mark, P.; Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954−9960.
(37) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. Dreiding - a Generic Force-Field for Molecular Simulations. *J. Phys. Chem.* **1990**, *94*, 8897−8909.
(38) Zhu, X.; Lopes, P. E. M.; MacKerell, A. D. Recent developments and applications of the CHARMM force fields. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 167−185.
(39) König, G.; Mei, Y.; Pickard, F. C.; Simmonett, A. C.; Miller, B. T.; Herbert, J. M.; Woodcock, H. L.; Brooks, B. R.; Shao, Y. Computation of Hydration Free Energies Using the Multiple Environment Single System Quantum Mechanical/Molecular Mechanical Method. *J. Chem. Theory Comput.* **2016**, *12*, 332−344.
(40) Sodt, A. J.; Mei, Y.; König, G.; Tao, P.; Steele, R. P.; Brooks, B. R.; Shao, Y. Multiple Environment Single System Quantum Mechanical/Molecular Mechanical (MESS-QM/MM) Calculations. 1. Estimation of Polarization Energies. *J. Phys. Chem. A* **2015**, *119*, 1511−1523.
(41) Lin, S. T.; Maiti, P. K.; Goddard, W. A. Two-Phase Thermodynamic Model for Efficient and Accurate Absolute Entropy of Water from Molecular Dynamics Simulations. *J. Phys. Chem. B* **2010**, *114*, 8191−8198.
(42) Pascal, T. A.; Lin, S. T.; Goddard, W. A. Thermodynamics of liquids: standard molar entropies and heat capacities of common solvents from 2PT molecular dynamics. *Phys. Chem. Chem. Phys.* **2011**, *13*, 169−181.
(43) Eyring, H.; Marchi, R. P. Significant structure theory of liquids. *J. Chem. Educ.* **1963**, *40*, 562.
(44) Marchi, R. P.; Eyring, H. Application of Significant Structure Theory to Water1. *J. Phys. Chem.* **1964**, *68*, 221−228.
(45) Yang, L.; Lin, S. T. Rapid prediction of solvation free energy and vapor pressure of liquid and solid from molecular dynamics simulation. *AIChE J.* **2015**, *61*, 2298−2306.
(46) Lai, P. K.; Hsieh, C. M.; Lin, S. T. Rapid determination of entropy and free energy of mixtures from molecular dynamics simulations with the two-phase thermodynamic model. *Phys. Chem. Chem. Phys.* **2012**, *14*, 15206−15213.
(47) Pascal, T. A.; Goddard, W. A. Interfacial Thermodynamics of Water and Six Other Liquid Solvents. *J. Phys. Chem. B* **2014**, *118*, 5943−5956.
(48) Jeon, J.; Kim, H.; Goddard, W. A.; Pascal, T. A.; Lee, G. I.; Kang, J. K. The Role of Confined Water in Ionic Liquid Electrolytes for Dye-Sensitized Solar Cells. *J. Phys. Chem. Lett.* **2012**, *3*, 556−559.
(49) Shin, H.; Pascal, T. A.; Goddard, W. A.; Kim, H. Scaled Effective Solvent Method for Predicting the Equilibrium Ensemble of Structures with Analysis of Thermodynamic Properties of Amorphous Polyethylene Glycol-Water Mixtures. *J. Phys. Chem. B* **2013**, *117*, 916−927.
(50) Weeks, J. D.; Chandler, D.; Andersen, H. C. Role of Repulsive Forces in Determining Equilibrium Structure of Simple Liquids. *J. Chem. Phys.* **1971**, *54*, 5237.
(51) Casanova, D.; Gusarov, S.; Kovalenko, A.; Ziegler, T. Evaluation of the SCF combination of KS-DFT and 3D-RISM-KH; Solvation effect on conformational equilibria, tautomerization energies, and activation barriers. *J. Chem. Theory Comput.* **2007**, *3*, 458−476.
(52) Li, X.-Y.; Wang, Q.-D.; Wang, J.-B.; Ma, J.-Y.; Fu, K.-X.; He, F.-C. Nonequilibrium solvation energy by means of constrained equilibrium thermodynamics and its application to self-exchange electron transfer reactions. *Phys. Chem. Chem. Phys.* **2010**, *12*, 1341−1350.
(53) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502−395520.
(54) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular-Dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.
(55) Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. Predicting small-molecule solvation free energies: An informal blind test for computational chemistry. *J. Med. Chem.* **2008**, *51*, 769−779.
(56) Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. Performance of SM8 on a test to predict small-molecule solvation free energies. *J. Phys. Chem. B* **2008**, *112*, 8651−8655.
(57) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(58) Hoover, W. G. Canonical Dynamics - Equilibrium Phase-Space Distributions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695−1697.
(59) Nose, S. A Unified Formulation of the Constant Temperature Molecular-Dynamics Methods. *J. Chem. Phys.* **1984**, *81*, 511−519.
(60) Price, D. J.; Brooks, C. L. A modified TIP3P water potential for simulation with Ewald summation. *J. Chem. Phys.* **2004**, *121*, 10096−10103.
(61) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(62) Ewald, P. P. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* **1921**, *369*, 253−287.
(63) Woutersen, S.; Emmerichs, U.; Bakker, H. J. Femtosecond Mid-IR Pump-Probe Spectroscopy of Liquid Water: Evidence for a Two-Component Structure. *Science* **1997**, *278*, 658−660.
(64) Liu, Y.; Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A. PAMAM Dendrimers Undergo pH Responsive Conformational Changes without Swelling. *J. Am. Chem. Soc.* **2009**, *131*, 2798−2799.
(65) Rezac, J.; Riley, K. E.; Hobza, P. S66: A Well-balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures. *J. Chem. Theory Comput.* **2011**, *7*, 2427−2438.
(66) Genheden, S.; Martinez, A. I. C.; Criddle, M. P.; Essex, J. W. Extensive all-atom Monte Carlo sampling and QM/MM corrections in the SAMPL4 hydration free energy challenge. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 187−200.
(67) Konig, G.; Pickard, F. C.; Mei, Y.; Brooks, B. R. Predicting hydration free energies with a hybrid QM/MM approach: an evaluation of implicit and explicit solvation models in SAMPL4. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 245−257.
(68) Reinisch, J.; Klamt, A. Prediction of free energies of hydration with COSMO-RS on the SAMPL4 data set. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 169−173.
(69) Ben-Amotz, D.; Underwood, R. Unraveling water's entropic mysteries: A unified view of nonpolar, polar, and ionic hydration. *Acc. Chem. Res.* **2008**, *41*, 957−967.

DOI: 10.1021/acs.jctc.6b00469
J. Chem. Theory Comput. 2016, 12, 5088−5099

(70) Honig, B.; Sharp, K.; Yang, A. S. Macroscopic Models of Aqueous-Solutions - Biological and Chemical Applications. *J. Phys. Chem.* **1993**, *97*, 1101−1109.
(71) Pellegrini, E.; Field, M. J. A generalized-born solvation model for macromolecular hybrid-potential calculations. *J. Phys. Chem. A* **2002**, *106*, 1316−1326.
(72) Rapp, C. S.; Friesner, R. A. Prediction of loop geometries using a generalized born model of solvation effects. *Proteins: Struct., Funct., Genet.* **1999**, *35*, 173−183.
(73) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free-Energies Using Macroscopic Solvent Models. *J. Phys. Chem.* **1994**, *98*, 1978−1988.
(74) Chocholoušová, J.; Feig, M. Balancing an accurate representation of the molecular surface in generalized born formalisms with integrator stability in molecular dynamics simulations. *J. Comput. Chem.* **2006**, *27*, 719−729.
(75) Gallicchio, E.; Kubo, M. M.; Levy, R. M. Enthalpy-entropy and cavity decomposition of alkane hydration free energies: Numerical results and implications for theories of hydrophobic solvation. *J. Phys. Chem. B* **2000**, *104*, 6271−6285.
(76) Harris, R. C.; Drake, J. A.; Pettitt, B. M. Multibody correlations in the hydrophobic solvation of glycine peptides. *J. Chem. Phys.* **2014**, *141*, 22D525.
(77) Harris, R. C.; Pettitt, B. M. Effects of geometry and chemistry on hydrophobic solvation. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 14681−14686.
(78) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Fixed-Charge Atomistic Simulations. *J. Chem. Theory Comput.* **2009**, *5*, 350−358.
(79) Ashbaugh, H. S.; Kaler, E. W.; Paulaitis, M. E. Hydration and conformational equilibria of simple hydrophobic and amphiphilic solutes. *Biophys. J.* **1998**, *75*, 755−768.
(80) Harris, R. C.; Pettitt, B. M. Examining the Assumptions Underlying Continuum-Solvent Models. *J. Chem. Theory Comput.* **2015**, *11*, 4593−4600.
(81) Ashbaugh, H. S.; Kaler, E. W.; Paulaitis, M. E. A "universal" surface area correlation for molecular hydrophobic phenomena. *J. Am. Chem. Soc.* **1999**, *121*, 9243−9244.
(82) Wallqvist, A.; Covell, D. G. Free-Energy Cost of Bending *n*-Dodecane in Aqueous Solution. Influence of the Hydrophobic Effect and Solvent Exposed Area. *J. Phys. Chem.* **1995**, *99*, 13118−13125.
(83) Vargaftik, N. B.; Volkov, B. N.; Voljak, L. D. International Tables of the Surface-Tension of Water. *J. Phys. Chem. Ref. Data* **1983**, *12*, 817−820.
(84) Huang, D. M.; Geissler, P. L.; Chandler, D. Scaling of hydrophobic solvation free energies. *J. Phys. Chem. B* **2001**, *105*, 6704−6709.
(85) Reiss, H.; Frisch, H. L.; Helfand, E.; Lebowitz, J. L. Aspects of the Statistical Thermodynamics of Real Fluids. *J. Chem. Phys.* **1960**, *32*, 119−124.
(86) Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature* **2005**, *437*, 640−647.
(87) Pan, A.; Biswas, T.; Rakshit, A. K.; Moulik, S. P. Enthalpy-Entropy Compensation (EEC) Effect: A Revisit. *J. Phys. Chem. B* **2015**, *119*, 15876−15884.

DOI: 10.1021/acs.jctc.6b00469
J. Chem. Theory Comput. 2016, 12, 5088−5099