THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

A "Stepping Stone" Approach for Obtaining Quantum Free Energies of Hydration

Chris Sampson,[†] Thomas Fox,[‡] Christofer S. Tautermann,[‡] Christopher Woods,[§] and Chris-Kriton Skylaris*[,†]

[†]School of Chemistry, University of Southampton, University Road, Southampton, Hampshire, SO17 1BJ, United Kingdom

[‡]Computational Chemistry, Lead Identification and Optimization Support, Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany

[§]School of Chemistry, University of Bristol, Cantocks Close, Bristol, Somerset, BS8 1TS, United Kingdom

ABSTRACT: We present a method which uses DFT (quantum, QM) calculations to improve free energies of binding computed with classical force fields (classical, MM). To overcome the incomplete overlap of configurational spaces between MM and QM, we use a hybrid Monte Carlo approach to generate quickly correct ensembles of structures of intermediate states between a MM and a QM/MM description, hence taking into account a great fraction of the electronic polarization of the quantum system, while being able to use thermodynamic integration to compute the free energy of transition between the MM and QM/MM. Then, we perform a final transition from QM/MM to full QM using a one-step free energy perturbation approach. By using QM/MM as a stepping stone toward the full QM description, we find very small convergence errors (<1 kJ/mol) in the transition to full QM. We apply this method to compute hydration free energies, and we obtain consistent improvements over the MM values for all molecules we used in this study. This approach requires large-scale DFT calculations as the full QM systems involved the ligands and all waters in their simulation cells, so the linear-scaling DFT code ONETEP was used for these calculations.

1. INTRODUCTION

Finding a computationally tractable method for computing accurate free energies of binding, typically between proteins and ligands, is one of the main challenges of computational drug optimization.[1−4] Typically, for large biologically relevant systems, force fields are used,[5−8] introducing dependency of the results on the parameters of the force field. The use of classical mechanics usually does not account for interactions which can only emerge from an *ab initio* quantum mechanical description, such as the exchange and the electronic charge transfer and polarization.[9]

A further great complexity of biomolecular free energy calculations is the requirement for extensive conformational sampling. Ideally, all conformational sampling would be performed using *ab initio* calculations, but this is far too computationally expensive to be feasible. A compromise between chemical accuracy and speed involves treating only part of the system (such as the binding pocket) with quantum mechanics and treating the rest with classical mechanics; this is known as the QM/MM approach. There are many applications of this method to calculate binding free energies.[10−12] QM/MM can require a large degree of expertise in the way the system should be partitioned into MM and QM regions, and setting up simulations for a nonexpert can be nontrivial. Additionally, if the QM region is chemically bound to the classical region, one needs to cleave chemical bonds and cap them with appropriate functional groups, for which chemically motivated approaches exist.[13]

A different strategy is to avoid direct conformational sampling using QM methods and instead define MM and QM ensembles and develop an approach for calculating the change in free energy between the two. This was first described by Warshel et al.[14] when he introduced an extended thermodynamic cycle which, for the calculation of free energies, has the form shown in Figure 1.

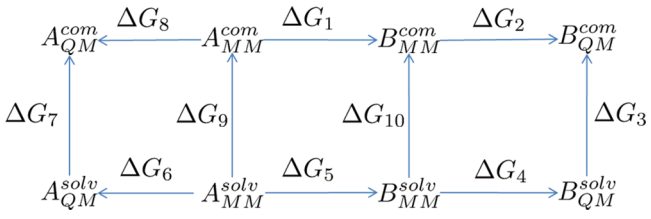

Figure 1. Extended thermodynamic cycle for computing free energies with QM methods.

Here, *A* and *B* are different molecular systems; commonly these are different ligands bound to the same protein. ΔG_3, ΔG_7, ΔG_9, and ΔG_{10} correspond to the standard (absolute) binding energy. Calculation of these values cannot be done directly. Computationally more tractable are alchemical mutations of one structure into another, in the complex

Received: February 17, 2015
Revised: May 6, 2015
Published: May 18, 2015

DOI: 10.1021/acs.jpcb.5b01625
J. Phys. Chem. B 2015, 119, 7030−7040

("com") and in solvent ("solv"), respectively (ΔG_1 and ΔG_5). In this cycle, ΔG_2, ΔG_4, ΔG_6, and ΔG_8 are the QM corrections to the classical free energies (ΔG_1 and ΔG_5).

As free energy is a state function, one can define the change in free energy in going from the classical to the quantum description. This has been applied, in the context of the cycle of Figure 1, by sampling structures using classical mechanics molecular dynamics, extracting regularly spaced uncorrelated snapshots, and perturbing to more expensive QM-based potentials.[15] Such methods, while alleviating some of the errors due to the classical force field, are still subject to the sampling issues associated with the classical, computationally cheaper potential. One possible way to apply this method is with the single-step free energy perturbation approach, i.e., using the structures generated from the classical MD simulation to calculate their fully quantum or QM/MM energy and then using the Zwanzig equation[16] (eq 1) to calculate the MM to QM change in free energy

$$\Delta G = -k_B T \ln\langle \exp[-\beta \Delta U_{QM-MM}]\rangle_{MM} \quad (1)$$

where k_B is the Boltzmann constant, T is the temperature, ΔU_{QM-MM} is the potential energy difference between the desired potential (in this case QM) and the sampling potential (in this case MM), β is $1/k_B T$, $\langle \cdots \rangle$ represents an ensemble average, and the subscript MM represents the potential used to obtain the structures.

The Zwanzig equation is derived under the assumption of using total potential energies and is exact in the limit of infinite sampling. However, if it is applied in this way for potentials that do not share a high degree of configurational space overlap, sampling issues occur;[17] e.g., structures that are unfavorable in the desired potential may be heavily sampled in the sampling potential and vice versa. These issues can be observed by gradually increasing the number of snapshots included within the Zwanzig equation. If large "jumps" occur in the free energy as the snapshots are increased, these are characteristic of poor sampling of the desired potential by the sampling potential,[18,19] as in these cases the discrepancy in the two potentials tends to create distinct snapshots for which the ΔU_{QM-MM} is large and negative affecting strongly the convergence of the Zwanzig equation, as it is an exponential average. If these "jumps" in free energy are suitably small, they can be smoothed out with more approximate approaches, such as the cumulant expansion of eq 1.[16,20−22] However, due to the large difference in magnitudes between the classical and quantum energies and the insufficient overlap between the classical and quantum configurational space, we observe a lack of convergence even when the cumulant expansion is applied. References 15 and 23 apply the equation by comparing between components of the interaction energy or the full interaction energies, rather than the total potential energy. Beierlein et al.[15] use the difference between the Coulombic contribution to the interaction energy, and Fox et al.[23] use the full quantum interaction energy where they perform density functional theory (DFT) calculations on the whole ligand−solvent system using the ONETEP linear-scaling DFT program.[24]

Interaction energies are defined by

$$U^{int} = U^{com} - U^{lig} - U^{host} \quad (2)$$

where the U^{com} is the potential energy of the entire system, e.g., a ligand bound to a protein and in the same example U^{host} would be the energy of the protein in the same geometry without the ligand present and U^{lig} is the energy of the ligand. U^{int} cancels out the intermolecular interactions within the host. We should note that intramolecular terms do not cancel out exactly in the QM description, as in this case pairwise additivity does not apply.

To overcome the difficulties associated with using classical force fields as a guiding potential for quantum mechanics due to the inexact overlap of configurational space,[25] Woods et al.[26] applied an acceptance criterion that allowed them to use MM to generate structures that were statistically correct for a QM/MM ensemble. The application of this method employs a Monte Carlo (MC) technique to generate MM structures and a Metropolis Hastings MC criterion to accept or reject these structures to the QM/MM ensemble. By adapting this method using hybrid Monte Carlo (HMC),[27] it is possible to minimize the number of quantum calculations while obtaining uncorrelated structures.

An alternative approach for generating a correct quantum ensemble would be to use the HMC technique[27] which is based on MD and therefore allows larger moves between acceptance tests and does not suffer random walk errors associated with standard Monte carlo techniques.[28,29] In this paper, we propose a HMC based approach for correcting classical free energies with quantum techniques. To do this, we use HMC to generate from MM an ensemble of QM/MM structures that is a much closer representation of the fully quantum ensemble. By doing this, the errors associated with sampling unfavorable structures are significantly lessened and also states intermediate between MM and QM/MM are generated which allow us to compute the MM to QM/MM change in free energy using thermodynamic integration (TI) which has stable convergence. We can then apply a single-step free energy perturbation from our generated QM/MM ensemble to the fully quantum ensemble. In order to validate our method, we calculate hydration free energies for five ligands. To our knowledge, this is the first time that HMC has been used in this manner for free energy calculations.

In section 2, we describe the theory and programs we have used including our three-step extended thermodynamic cycle that goes from MM to QM/MM to full QM. In section 3, we present and discuss our results on the hydration free energies of ligands as obtained with MM, QM/MM, and full QM, and in section 4, we finish with our conclusions.

2. METHODS

Our method first generates an ensemble of structures closer to the fully quantum ensemble by applying the hybrid Monte Carlo method using classical molecular dynamics to sample configurations but accepting a QM/MM ensemble. The QM/MM ensemble energy is calculated as the energy of the MM system but with the Coulombic component of its interaction energy replaced by the equivalent interaction energy in the quantum description. This takes into account the electronic polarization of the ligand as a result of the surrounding solvent. Once an ensemble of QM/MM structures has been generated, then a single-step perturbation approach can be applied to calculate the free energy of mutation from the QM/MM to the full QM ensemble.

2.1. Theoretical Details. *2.1.1. The Hybrid Monte Carlo Method.* To generate an ensemble of structures that corresponds to a "target" potential from a "guiding" potential used for sampling, the traditional hybrid Monte Carlo (HMC) method can be used.[27] The difference between the target and the guide states can involve differences in the ensembles. For

DOI: 10.1021/acs.jpcb.5b01625
J. Phys. Chem. B 2015, 119, 7030−7040

example, the guide can be an MM NVE ensemble and the target can be a QM NVT ensemble. By using the HMC method, the detailed balance condition (eq 3) is satisfied

$$\rho(\mathbf{R})\pi(\mathbf{R} \rightarrow \mathbf{R}') = \rho(\mathbf{R}')\pi(\mathbf{R}' \rightarrow \mathbf{R}) \tag{3}$$

where $\rho(\mathbf{R})$ is the probability of occupying a certain configuration $\mathbf{R}$ in the target ensemble. From here onward, this will be referred to as state probability. $\pi(\mathbf{R} \rightarrow \mathbf{R}')$ is the transition probability shown in eq 4, which consists of the trial probability and the acceptance probability

$$\pi(\mathbf{R} \rightarrow \mathbf{R}') = \pi_{\text{acc}}(\mathbf{R} \rightarrow \mathbf{R}')t(\mathbf{R} \rightarrow \mathbf{R}') \tag{4}$$

where $t(\mathbf{R} \rightarrow \mathbf{R}')$ is the trial probability, which is the probability of moving to a new configuration and $\pi_{\text{acc}}(\mathbf{R} \rightarrow \mathbf{R}')$ is the acceptance probability of going from $\mathbf{R}$ to $\mathbf{R}'$ in the target potential, given by a Metropolis–Hastings criterion, and is given by eq 5.

$$\pi_{\text{acc}}(\mathbf{R} \rightarrow \mathbf{R}') = \min\left\{1, \frac{\rho(\mathbf{R}')t(\mathbf{R}' \rightarrow \mathbf{R})}{\rho(\mathbf{R})t(\mathbf{R} \rightarrow \mathbf{R}')}\right\} \tag{5}$$

The generation of new configurations in the HMC method is performed by an underlying MD simulation. The advantages of using the HMC method are that the sampling errors associated with the target potential are avoided and the configurational differences between $\mathbf{R}$ and $\mathbf{R}'$ can be large.[28] To satisfy a detailed balance, the MD simulation in the guiding potential must be reversible; therefore, we have chosen to run within the microcanonical ensemble. The acceptance probability shown in eq 5 requires inputs at points $\mathbf{R}$ and $\mathbf{R}'$ which are the starting and end points of an MD simulation, respectively. The state probability is the probability of being in a certain conformation in the target potential, which for both the original configuration and the final configuration is, according to the rules of statistical mechanics, given by the Boltzmann distribution

$$\rho(\mathbf{R}) = \frac{\exp(-\beta U(\mathbf{R}))}{Z(N, V, T)} \tag{6}$$

where $Z(N, V, T)$ is the configurational partition function and U is the energy of $\mathbf{R}$ in the target potential.

In addition to the state probability, the trial probability (eq 7) is required in eq 5. This is the probability of transitioning from one structure to another, e.g., from $\mathbf{R}$ to $\mathbf{R}'$. In theory, MD simulations are deterministic, i.e., if a simulation is given the same structure and velocities and run for the same length of time, the final structure will always be the same. As such, the trial probability is directly proportional to the momenta given at the start of the simulation. In order to create a Markov chain of structures, random momenta are provided at the start of each MD simulation. These random momenta are taken from a Gaussian distribution using the Marsaglia polar method.[30] The Gaussian is used so the Boltzmann distribution of kinetic energies is satisfied. Additionally, by selecting momenta from a Gaussian, we can set the temperature, thus providing a thermostat.

$$t(\mathbf{R} \rightarrow \mathbf{R}') \propto \prod_{i=1}^{n} \frac{1}{\sqrt{2m_i\pi k_B T}} \exp\left[-\beta \frac{\mathbf{p}_i^2}{2m_i}\right] \tag{7}$$

To derive a workable form of the acceptance criterion, we place eqs 6 and 7 into eq 5. By noticing that eq 7 is the functional form of kinetic energy ($K(\mathbf{P})$, where $\mathbf{P}$ are the momenta), we obtain

$$\begin{aligned}\pi_{\text{acc}}(\mathbf{R} \rightarrow \mathbf{R}') &= \min\left\{1, \frac{\exp(-\beta U(\mathbf{R}'))\exp(-\beta K(\mathbf{P}'))}{\exp(-\beta U(\mathbf{R}))\exp(-\beta K(\mathbf{P}))}\right\} \\ &= \min\left\{1, \frac{\exp(-\beta H(\mathbf{R}', \mathbf{P}'))}{\exp(-\beta H(\mathbf{R}, \mathbf{P}))}\right\}\end{aligned} \tag{8}$$

The method presented here aims to mutate a solvent–ligand complex from its MM representation to a "quantum corrected" representation where the classical electrostatic interactions have been replaced by interactions from a QM/MM calculation. The acceptance criterion is then

$$\pi_{\text{acc}}(\mathbf{R} \rightarrow \mathbf{R}') = \min\left\{1, \frac{\exp(-\beta U_{\text{MM+QM}^{\text{int}}_{\text{Coul}}}(\mathbf{R}'))\exp(-\beta K(\mathbf{P}'))}{\exp(-\beta U_{\text{MM+QM}^{\text{int}}_{\text{Coul}}}(\mathbf{R}))\exp(-\beta K(\mathbf{P}))}\right\} \tag{9}$$

where the target potential energy is calculated by

$$\begin{aligned}U_{\text{MM+QM}^{\text{int}}_{\text{Coul}}} &= U^{\text{com}}_{\text{MM}} - [U^{\text{com}}_{\text{MM}_{\text{Coul}}} - U^{\text{host}}_{\text{MM}_{\text{Coul}}} \\ &\quad - U^{\text{lig}}_{\text{MM}_{\text{Coul}}}] + [U^{\text{com}}_{\text{QM/MM}} - U^{\text{host}}_{\text{QM/MM}} \\ &\quad - U^{\text{lig}}_{\text{QM/MM}}]\end{aligned} \tag{10}$$

$$= U^{\text{com}}_{\text{MM}} - U^{\text{int}}_{\text{MM}_{\text{Coul}}} + U^{\text{int}}_{\text{QM}_{\text{Coul}}} \tag{11}$$

The above equation uses the classical potential energy ($U^{\text{com}}_{\text{MM}}$) for the whole complex and subtracts from it the electrostatic (Coulomb, "Coul") contribution of the classical interaction energy ($U^{\text{int}}_{\text{MM}_{\text{Coul}}}$). This interaction energy is replaced by the interaction energy between the QM/MM ligand ($U^{\text{lig}}_{\text{QM/MM}}$) and the host ($U^{\text{host}}_{\text{QM/MM}}$) from the QM/MM calculation (in this work, the host is the solvent). The QM/MM description used here is the quantum ligand surrounded by classical point charges. The dispersion (Lennard-Jones) part of the interaction energy is not replaced—it still comes from the MM calculation.

2.1.2. Transitioning between MM and QM/MM. In order to make the transition from U_{MM} to $U_{\text{MM+QM}^{\text{int}}_{\text{Coul}}}$ smoothly, intermediate $\lambda_{\text{QM/MM}}$ steps can be introduced. The introduction of these steps is trivial, and can be performed by changing the $U_{\text{MM+QM}^{\text{int}}_{\text{Coul}}}(\mathbf{R})$ term of eq 9 with the following

$$\begin{aligned}U_{\text{MM+QM}^{\text{int}}_{\text{Coul}}}(\mathbf{R}; \lambda_{\text{QM/MM}}) &= (1 - \lambda_{\text{QM/MM}})U_{\text{MM}} + (\lambda_{\text{QM/MM}})U_{\text{MM+QM}^{\text{int}}_{\text{Coul}}}\end{aligned} \tag{12}$$

HMC is run for all lambda states between the MM and QM/MM, such that at $\lambda_{\text{QM/MM}} = 0$ we are accepting to a classical canonical ensemble and at $\lambda_{\text{QM/MM}} = 1$ to the QM/MM corrected canonical ensemble. The difference for each lambda value is in the MC acceptance criterion which influences the progress of the MD simulations accordingly. The free energy between the MM and QM/MM is then obtained using thermodynamic integration (TI)

DOI: 10.1021/acs.jpcb.5b01625
J. Phys. Chem. B 2015, 119, 7030–7040

$$\Delta G = \int_0^1 d\lambda_{QM/MM} \left\langle \frac{\partial U(\lambda_{QM/MM})}{\partial \lambda_{QM/MM}} \right\rangle_{\lambda_{QM/MM}}$$

$$\frac{\partial U(\lambda_{QM/MM})}{\partial \lambda_{QM/MM}} = U_{MM+QM^{int}_{Coul}} - U_{MM}$$

$$\Delta G = \int_0^1 d\lambda_{QM/MM} \langle U_{MM+QM^{int}_{Coul}} - U_{MM} \rangle_{\lambda_{QM/MM}} \quad (13)$$

$$= \int_0^1 d\lambda_{QM/MM} \langle U^{int}_{QM_{Coul}} - U^{int}_{MM_{Coul}} \rangle_{\lambda_{QM/MM}} \quad (14)$$

Each lambda window is built up by running classical MD simulations and applying the acceptance test shown in eq 12. To clarify, no MD simulation that involves a mix of MM and QM/MM is needed.

2.1.3. Transitioning to the Full QM. We expect that the QM/MM ensemble is a much closer representation of the QM ensemble, so the transition from the QM/MM to QM should be possible with a method which is simpler than the HMC method. Therefore, to perform this perturbation, the energy differences $U_{QM} - U_{MM+QM^{int}_{Coul}}$ are calculated and then the Zwanzig equation is used to calculate the free energy change from QM/MM to full QM, as shown in eq 1.

We have also computed this free energy by using the interaction energy, as defined in eq 2, of the full quantum system, in place of the total energy U_{QM}. The full QM interaction energy in DFT contains also the attractive component of dispersion interactions which are given by various types of empirical correction.[31,32] We have therefore also investigated the effect of changing this empirical dispersion correction between different dispersion methods.

As we are doing the transition to full QM via a QM/MM state, our full thermodynamic cycle has three steps, as shown in Figure 2. Figure 3 shows the whole correction process.

2.2. Computational Details. As to our knowledge this is the first time such a three-step approach is used, our aim in this paper is to validate it in the calculation of hydration free energies, as preparation for application to more challenging host−ligand systems in the future. We applied our method to the calculation of hydration free energies for ethanol, ethane, ethylene glycol, dimethyl ether, and propane, each in a simulation cell with explicit waters. These molecules were chosen due to their different degrees of polarity, ranging from nonpolar (hydrophobic) such as ethane and propane to highly polar (hydrophilic) such as ethylene glycol.

2.2.1. Classical Simulation Details. The charges for the solute were obtained from AM1-BCC calculations using Antechamber;[33] the force field parameters were taken from the GAFF force field.[34] The water model used throughout was the TIP3P model.[35] Electrostatics were treated with PME, and a cutoff of 8 Å was used for van der Waals interactions. All classical MD simulations were performed within the double precision version of Gromacs v4.6.5,[36] to allow us to use the velocity Verlet integrator, as access to kinetic energies is required for the HMC method. The double precision was necessary to ensure good energy conservation in the NVE simulations.

Each ligand was solvated with 450 waters, and the whole system was equilibrated using the following procedure. Structures were initially minimized with the steepest descent algorithm for 5000 steps. Following this, a 500 ps simulation in the canonical ensemble was run using a time step of 1 fs and the Berendsen thermostat was used to heat the system, raising the temperature linearly from 100 to 300 K. Finally, a 1 ns isothermal−isobaric simulation was performed using the Berendsen barostat. After equilibration, the box size for each ligand was around 24 Å^3.

Using the NPT equilibrated structures, further equilibration was then applied using the HMC method where structures were accepted from the NVE ensemble into a classical NVT ensemble for 100 HMC steps. As the MD simulations used within the HMC method were run in the microcanonical ensemble, we expect to have an acceptance rate of 100% as the total Hamiltonian energy should be constant. However, due to fluctuations within the total energy during the NVE simulation, this is commonly not the case. To minimize these fluctuations, a time step of 0.25 fs was used. The length of these MD simulations was determined by tests described in section 3.

Classical relative free energies between systems were calculated using TI. Seventeen classical λ_{MM} windows were used, mutating the charges and VdW interactions at the same time using soft-core potentials. The classical λ_{MM} windows had the values 0.0, 0.002, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99, 0.998, and 1.0. The small λ_{MM} differences at the end points are to tackle discontinuities. Each classical λ_{MM} window underwent an equilibration procedure which is the same as that mentioned above, with the exception that the NPT simulation to equilibrate the box size was 500 ps. Following this, a production simulation of 2 ns was run. Errors associated with the TI free energies were calculated by hysteresis, and the free energies are the average of the forward and reverse calculations.

2.2.2. Quantum Simulation Details. For our DFT program, we have used ONETEP.[24] The ONETEP program is a linear-scaling DFT code that has been developed for use on parallel computers.[37] ONETEP combines linear scaling with accuracy comparable to conventional cubic-scaling plane-wave methods, which provide an unbiased and systematically improvable approach to DFT calculations. Its novel and highly efficient algorithms allow calculations on systems containing tens of thousands of atoms.[38] ONETEP is based on a reformulation of DFT in terms of the one-particle density matrix. The density matrix in terms of Kohn−Sham orbitals is

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{n=0}^{\infty} f_n \psi_n(\mathbf{r}) \psi_n'(\mathbf{r}') \quad (15)$$

where f_n is the occupancy and $\psi_n(\mathbf{r})$ are the Kohn−Sham orbitals. In ONETEP, the density matrix is represented as

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{\alpha} \sum_{\beta} \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi_\beta^*(\mathbf{r}') \quad (16)$$

where $\phi_\alpha(\mathbf{r})$ are localized nonorthogonal generalized Wannier functions (NGWFs)[39] and $K^{\alpha\beta}$, which is called the density kernel, is the representation of f_n in the duals of these functions. Linear scaling is achieved by truncation of the density kernel, which decays exponentially for materials with a band gap, and by enforcing strict localization of the NGWFs onto atomic regions. In ONETEP, as well as optimizing the density kernel, the NGWFs are also optimized, subject to a localization constraint. Optimizing the NGWFs *in situ* allows for a minimum number of NGWFs to be used while still achieving plane-wave accuracy. The NGWFs are expanded in a basis set of periodic sinc (psinc) functions,[40] which are equivalent to a

DOI: 10.1021/acs.jpcb.5b01625
J. Phys. Chem. B 2015, 119, 7030−7040

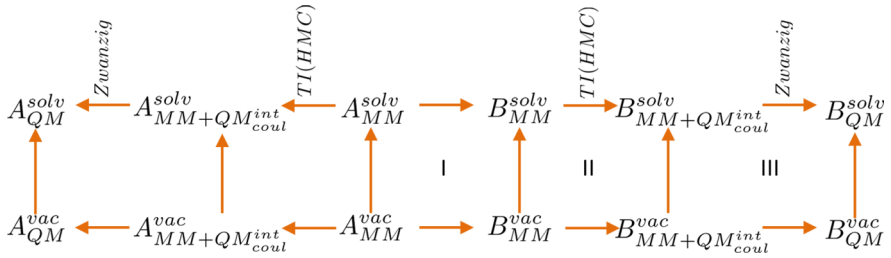

Figure 2. Three-step thermodynamic cycle. Cycle I shows the classical mutation from ligand A to B, cycle II describes the transition from MM to QM/MM, and finally cycle III shows the transition from QM/MM to full QM.

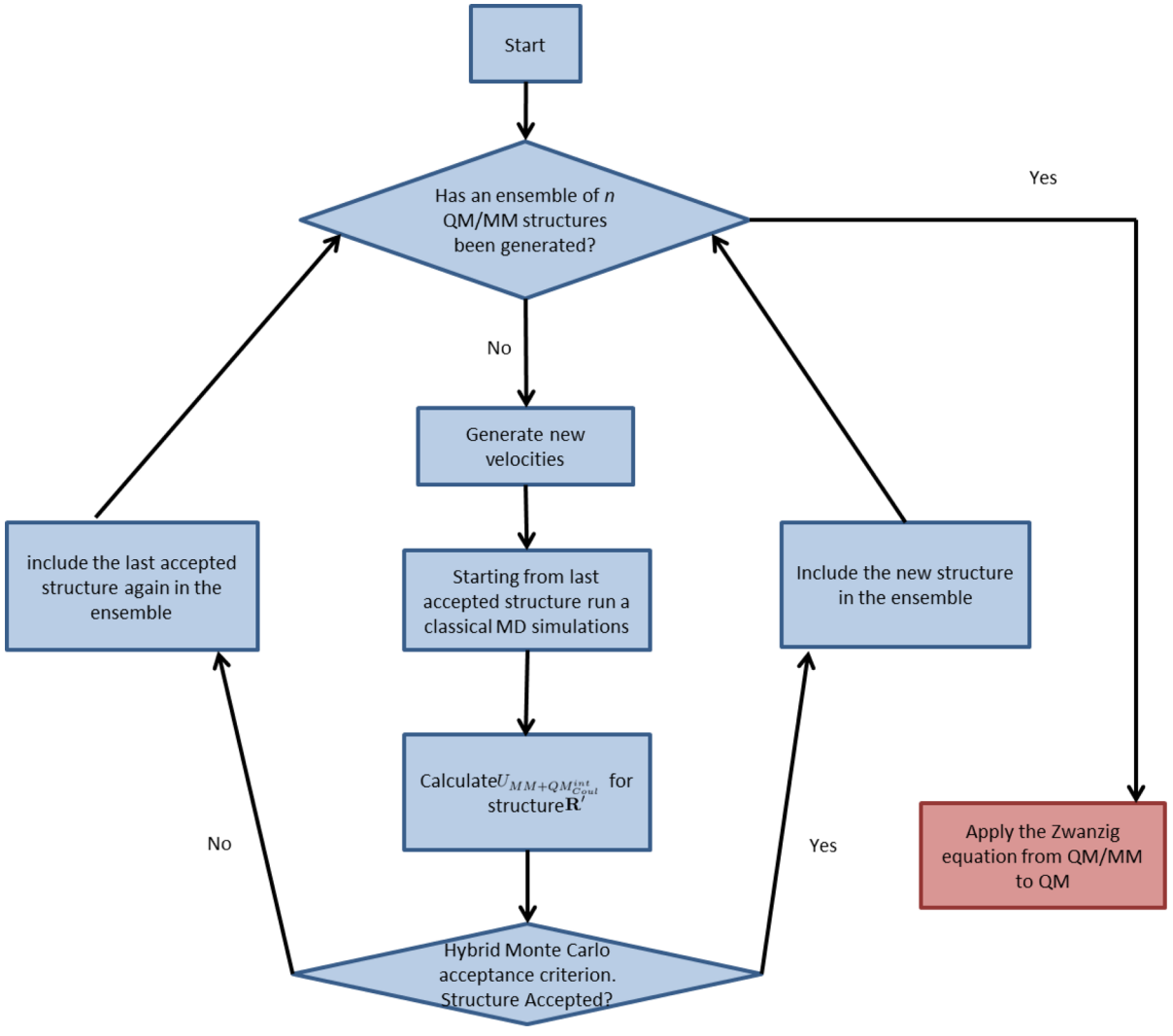

Figure 3. Flowchart for computing steps 2 and 3 of the three-step thermodynamic cycle of Figure 2.

plane-wave basis, as they are related by a unitary transformation. Using a plane-wave basis set allows the accuracy to be improved by varying a single parameter, equivalent to the kinetic energy cutoff in conventional plane-wave DFT codes. The psinc basis set provides a uniform description of space, meaning that ONETEP does not suffer from basis set superposition error.[41]

Following the classical 100-step equilibration within the HMC method, we switched from going from NVE (MM) to NVT (MM) to going from NVE (MM) to NVT (QM/MM). The first 100 steps of this process were also taken as equilibration where the QM/MM energy was computed according to eq 10. After this, 500 steps were run and counted as production, for the transition to QM/MM. This process was repeated, accepting different values of $\lambda_{QM/MM}$. We used three values, $\lambda_{QM/MM}$ = 0, 0.5, 1.

Each quantum simulation was performed using ONETEP.[24] Four NGWFs were used on heavy atoms and one on hydrogen, all with a radius of 8.0 a_0. Calculations used a psinc kinetic energy cutoff of 800 eV. Embedding charges[42] were used to represent water within the calculations, so only the ligand was fully represented by QM. Only the electrostatics were corrected by quantum mechanics, so dispersion was included at the classical level, as in eq 10.

DOI: 10.1021/acs.jpcb.5b01625
J. Phys. Chem. B 2015, 119, 7030−7040

The fully quantum calculations for cycle III of the thermodynamic cycle of Figure 2 were obtained by restoring the embedding charges to explicit atoms, keeping all other parameters the same with the exception of the dispersion component which, within the total energy perturbation, was treated by a damped London potential[31] using the "Elstner" method.[43] Within the interaction energy perturbation approach, the "Elstner",[43] "Grimme D2",[32] "Grimme D3",[44] and the dispersion component of the force field were all tested. When the force field dispersion was used within the correction, this was extracted from the classical simulation and used within the QM interaction energy.

2.3. Determining the Length of the MD Simulations Used within the HMC Method. In order to optimize the length of the MD simulations required to run the HMC method with a reasonable acceptance rate, the MD runs were increased progressively from 0.1 ps to 1 ns. This was applied to ethanol in 450 waters and was repeated for two runs to ensure convergence. Each HMC run consisted of 500 production steps following the equilibration procedure described in subsection 2.2.2. The results can be seen in Table 1. "Energy" refers to the

Table 1. Determining the MD Simulation Length within the HMC Method for Ethanol in 450 Waters[a]

	energy (kJ/mol)/acceptance (%)	
HMC step size (ps)	run 1	run 2
0.1	−16.08/14.0	−16.90/17.5
1	−15.46/16.2	−16.56/11.5
10	−16.01/13.6	−14.99/13.3
100	−16.78/8.6	−14.27/16.6
1000	−13.64/21.6	−16.21/11.2

[a]Energy is the average of $(U^{int}_{QM_{Coul}} - U^{int}_{MM_{Coul}})$ at $\lambda_{QM/MM} = 1$.

average energy differences $(U^{int}_{QM_{Coul}} - U^{int}_{MM_{Coul}})$ at $\lambda_{QM/MM} = 1$. These energy differences are required for TI (eq 14). By examining the energy differences at $\lambda_{QM/MM} = 1$ where the acceptance will be lowest, we can confirm the acceptance at other windows will also be sufficient.

The largest difference between two HMC runs shown is 2.51 kJ/mol, which is around thermal error (k_BT, 2.5 kJ/mol). Similarly, there is only a small amount of error between simulations of different lengths. This suggests that the energy difference is essentially independent of the length of the MD simulation used for this system. In order to ensure uncorrelated snapshots while keeping the length of CPU time required for each MD manageable, 100 ps was chosen as the HMC step size, for all our subsequent HMC simulations. The 1000 ps simulations show no substantial improvement on the 100 ps run, while taking an order of magnitude longer to run.

3. RESULTS AND DISCUSSION

3.1. Classical TI. The classical free energy of mutation (relative hydration free energies) between our ligands was calculated using ethanol as the reference. The values are shown in Table 4 in the "classical" column. All the experimental values were obtained from ref 45 with the exception of that for ethylene glycol, which was taken from ref .

The mutation to propane shows the largest hysteresis which is 1.2 kJ/mol, which can be attributed either to the force field or to the introduction of two additional atoms, whereas all other mutations involved either one or no additional atoms. All calculated free energies were within thermal error (2.5 kJ/mol) from the experimental free energies, with the exception of ethanol to ethane which, while being a well converged mutation, is 7.2 kJ/mol away from the experimental value, pointing to force field limitations.

3.2. Calculating the Free Energy of Mutation from MM to QM/MM. The resulting $dV/d\lambda_{QM/MM}$ values as a function of $\lambda_{QM/MM}$ value from the application of the HMC method to all test systems are shown in Figure 4. Three runs were performed for each ligand, and in each case, they are consistent and converge to the same value.

The largest differences between the same $\lambda_{QM/MM}$ window are shown by ethanol and ethylene glycol. In the case of ethanol, the largest difference of 2.5 kJ/mol between the three runs is observed at $\lambda_{QM/MM} = 1$ and the difference for ethylene glycol is much larger at 5.8 kJ/mol again at $\lambda_{QM/MM} = 1$. This can be explained by the presence of the hydroxyl groups forming much stronger bonds with the surrounding water than any of the other ligands used, leading to low acceptance. Indeed, this is reflected by the values shown within the graph. Ethane shows the smallest energy difference in going from $\lambda_{QM/MM} = 0$ to $\lambda_{QM/MM} = 1$. This can be expected, as it is the most apolar, meaning that the quantum Coulombic contribution to the interaction energy was very close to the classical equivalent. Following this trend, propane is the next apolar, and thus has the next smallest energy difference, then dimethyl ether, followed by ethanol, and then ethylene glycol.

This trend is also reflected within the acceptance ratios (shown in brackets within Figure 4). Again, the lowest acceptance is found when applying the HMC method to ethylene glycol, with an acceptance below 10%. This low acceptance is caused by the large mismatch between the energy differences, which shows an inconsistency between the QM and MM electrostatic interaction energies. This fluctuation of energy differences could be explained by the strong electronic polarization of polar ligands. Polarization is implicitly included within the AMBER force field and explicitly within the QM/MM description. Implicit polarization may be inaccurate in a chemical environment that varies from that of water (e.g., a binding pocket), but the force field nevertheless shows excellent correlation with experimental hydration free energies ("classical" column, Table 4).

In order to assess the polarization effect on these ligands, the dipole moments of structures that were accepted to the QM/MM ensemble for dimethyl ether and ethylene glycol were calculated both classically and within ONETEP. The classical dipole moment was calculated once per structure, given that the force field is not polarizable, so the dipole does not change between the vacuum and solvent, but the quantum dipole moment was calculated twice, once in the solvent (embedding point charges) and once in the vacuum. The results are shown in Table 2 for five structures for each molecule that span the range of the dipole moments obtained by each calculation method.

The classical dipole moment for dimethyl ether is fairly constant between the five sampled snapshots, whereas the dipole moments for ethylene glycol fluctuate. The discrepancy between the classical dipole moments and those of the QM/MM system in solvent is markedly larger (0.2−0.4 eÅ) for the ethylene glycol than for dimethyl ether (discrepancies of less than 0.1 eÅ). In fact, for dimethyl ether, two of the classical dipole moments correctly match the quantum dipole moments, and with the exception of structures 4 and 5, the quantum

DOI: 10.1021/acs.jpcb.5b01625
J. Phys. Chem. B 2015, 119, 7030−7040

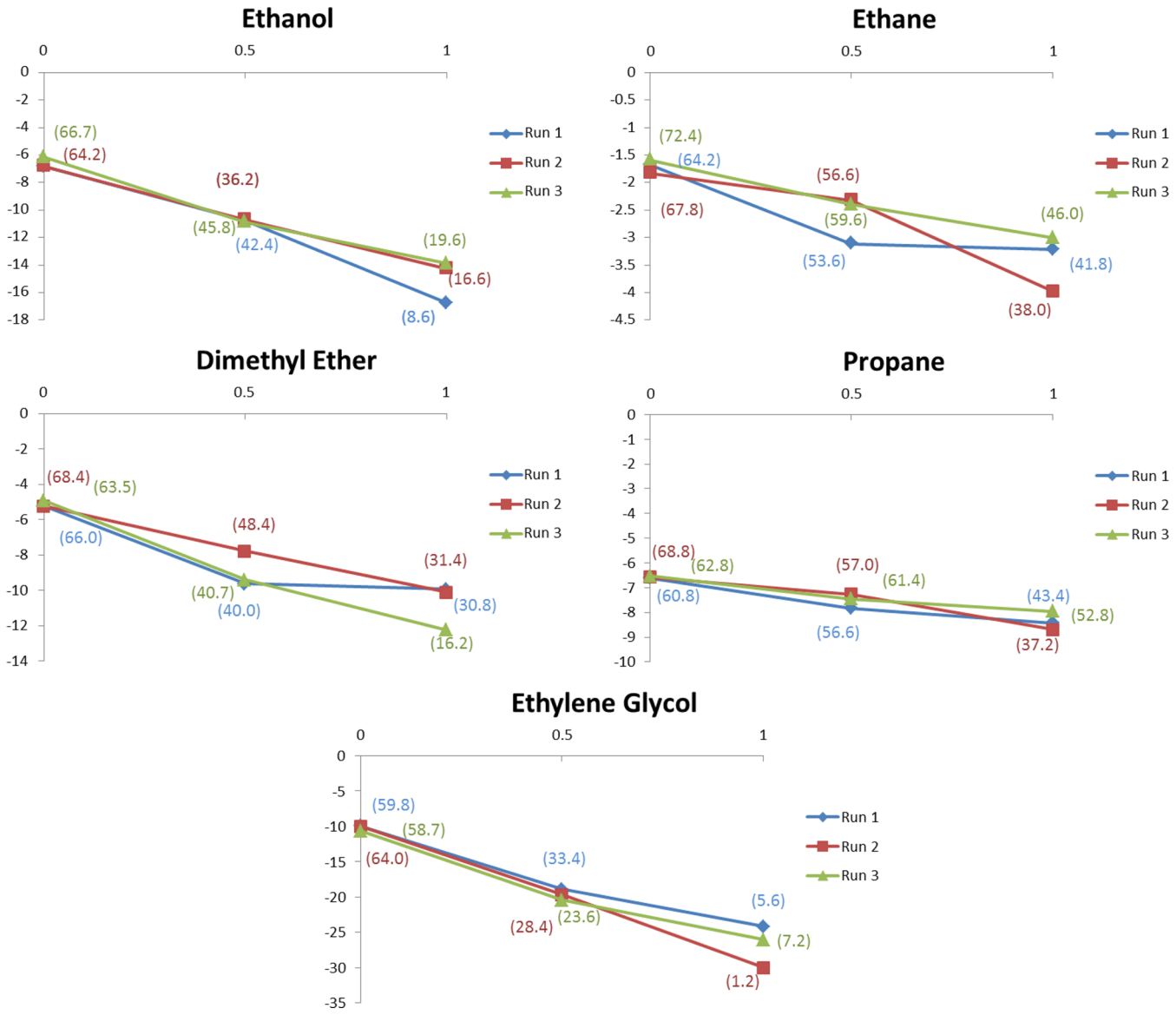

Figure 4. PMFs for moving between U_{MM} and $U_{MM+QM_{Coul}^{int}}$. Each point shows the average energy difference required for eq 14 in kJ/mol. These energy differences were calculated for three $\lambda_{QM/MM}$ windows for each system. The value in brackets shows the acceptance.

Table 2. Magnitude of the Dipole Moments in eÅ for Five Example Structures of Ethylene Glycol and Dimethyl Ether Accepted into the QM/MM Ensemble

ligand	structure	QM/MM solvated	QM/MM vacuum	classical
ethylene glycol	1	0.457	0.340	0.313
	2	0.635	0.530	0.579
	3	0.725	0.566	0.515
	4	0.720	0.492	0.641
	5	0.895	0.735	0.460
dimethyl ether	1	0.237	0.121	0.338
	2	0.304	0.218	0.378
	3	0.300	0.230	0.371
	4	0.318	0.229	0.320
	5	0.331	0.220	0.329

dipole moments for dimethyl ether within solvent match the experimental value of 0.271 eÅ extremely well, whereas classical mechanics overestimates it. Another interesting comparison between the methods can be made with a TIP3P water molecule, where the quantum dipole moment in a vacuum is calculated to be 0.380 eÅ and the classical dipole moment is 0.489 eÅ. The experimental value is 0.385 eÅ, showing again that the quantum value is extremely close and the classical value, again, is much less accurate.

These results highlight the fact that only by using quantum corrections we can account for the explicit polarization, as the force field can at best describe it in an implicit (average) manner, and indicate that the lower HMC acceptance for ethylene glycol may be a consequence of the large discrepancy in polarization between the classical and quantum descriptions.

Integrating each energy vs $\lambda_{QM/MM}$ value curve gives the free energy, in accordance with eq 14. These free energy values are shown in Table 3 for each of the three runs.

The magnitude of these free energies again follows the trend of the polarity of the molecules. The differences show the convergence between the free energies. In every case, convergence has been achieved. The largest difference is present for ethylene glycol of 1.86 kJ/mol; however, this value

DOI: 10.1021/acs.jpcb.5b01625
J. Phys. Chem. B 2015, 119, 7030−7040

Table 3. Free Energy of Changing from U_{MM} to $U_{MM+QM^{int}_{Coul}}$ [a]

	run 1	run 2	run 3	difference
ethanol	−11.29	−10.62	−10.45	0.84
ethane	−2.78	−2.61	−2.35	0.43
ethylene glycol	−17.97	−19.82	−19.38	1.86
dimethyl ether	−8.60	−7.72	−8.99	1.27
propane	−7.67	−7.45	−7.35	0.32

[a]All values shown are in kJ/mol. The difference is the largest difference between the three runs.

is still within thermal error (2.5 kJ/mol). By combining these free energies with the classical free energies shown in Table 4 under the "classical" column, according to cycles I and II of the three-step thermodynamic cycle (Figure 2), we obtain the QM/MM corrected free energies shown in Table 4 under the MM +QM^{int}_{Coul} corrected column.

The addition of the QM correction due to the MM to QM/MM transition to the classical relative free energies between ethanol and ethane shows a definite improvement of the free energy which was badly underestimated by the force field. The corrections for ethylene glycol and propane move the free energies in the direction of improvement, although they overlook the experimental values. The free energy for dimethyl ether is shifted further away from the experimental value, as the classical free energy was extremely close already. These results are encouraging and demonstrate the effect of QM/MM derived polarization on the free energies as an intermediate step toward the final stage where we will introduce the full QM description.

3.3. Calculating the Free Energy of Mutation from QM/MM to QM. Once a QM/MM ensemble of 500 structures was built up, the next step was the perturbation to the fully quantum system (see cycle II in Figure 2). This was performed in two ways: first by using the difference in total potential energies (eq 17) and second by using the difference in interaction energies (eq 18). In both cases, this perturbation was performed by using the Zwanzig equation (eq 1).

$$\Delta U^{total} = U_{QM} - U_{MM+QM^{int}_{Coul}} \quad (17)$$

$$\Delta U^{int} = U^{int}_{QM} - U^{int}_{MM+QM^{int}_{Coul}} \quad (18)$$

where U^{int}_{QM} is the interaction energy at the full QM level and $U^{int}_{MM+QM^{int}_{Coul}}$ is the interaction energy at the QM/MM level.

3.4. Total Energy Perturbation. The final calculated free energy when using total energies (eq 17) shows no consistency between the two runs which demonstrates lack of convergence of the exponential average of the Zwanzig equation. This is demonstrated for three of the ligands in the left panel of Figure 5. Figure 5 shows the free energy as a function of the number of snapshots, i.e., a running exponential average as the number of snapshots is increased. These graphs show that the free energy is affected dramatically by individual snapshots. The graphs show a large variation between the two runs for each ligand. For example, for ethylene glycol, a difference of 35.17 kJ/mol is present after 500 snapshots, while for dimethyl ether the difference between the runs is 0.33 kJ/mol. The small difference for dimethyl ether suggests convergence; however, if Figure 5 is examined, it is clear that large "jumps" in the free energy are present and there is no guarantee that they will not affect the results if we were to run more than 500 snapshots. This is described as "sawtooth" sampling and is characteristic of a lack of overlap between the configuration space.[18] This observation has also been made when moving between a purely classical ensemble and a quantum ensemble by Cave-Ayland et al.[47] They explain that this lack of configuration space overlap is due to differences in the intramolecular degrees of freedom. Therefore, by excluding all intramolecular terms (i.e., using interaction-energy-based corrections), convergence can be achieved. Indeed, this is what is observed here and will be covered in more detail in the next section.

To complete the three-step thermodynamic cycle when using total energies, an additional calculation must be performed of the ligand in the vacuum. This is displayed in the lower half of cycles II and III in Figure 2. However, the initial MM+QM^{int}_{Coul} correction required for the ligand in the vacuum (cycle II) would essentially be a standard classical MD, as the interaction energy in the vacuum is zero, and with the system size being relatively small, the fluctuation in the total Hamiltonian energy would be small, leading to a high acceptance. The ligand energy differences required for the mutation from MM to QM/MM in the Zwanzig equation (cycle III) were taken from values already calculated within the top section of cycle II. For each ligand, the calculated free energies converge to less than 0.2 kJ/mol between the two runs. This result combined with the fact that our water molecules are rigid leads to the conclusion that the lack of overlap between the configuration space should be attributed to the intermolecular interactions between the water molecules. While this appears to be the case for the small rigid ligands we use in this study for larger ligands, with more

Table 4. Relative Hydration Free Energies Using the Experimental Value for Ethanol (−20.92 kJ/mol) as a Reference[45] [a]

	free energy (kJ/mol)						
ligand	experimental	classical	MM+QM^{int}_{Coul} corrected	Elstner	force field disp	Grimme D2	Grimme D3
ethane	7.66[45]	0.53 ± 0.23	8.71 ± 1.27	10.31 ± 0.18	7.52 ± 0.54	11.07 ± 0.58	10.79 ± 0.67
ethylene glycol	−38.91[46]	−35.91 ± 0.20	−44.18 ± 2.70	−43.59 ± 1.52	−42.61 ± 3.20	−44.09 ± 1.67	−44.14 ± 1.60
dimethyl ether	−7.99[45]	−7.93 ± 0.02	−5.58 ± 2.11	−5.83 ± 0.26	−7.76 ± 0.76	−5.17 ± 0.54	−5.39 ± 0.42
propane	8.20[45]	6.93 ± 1.18	10.22 ± 1.16	13.74 ± 0.16	6.47 ± 0.50	14.56 ± 0.74	14.06 ± 0.54
RMS error		3.93	3.11	4.04	2.05	4.66	4.42
max error		7.15	5.27	5.54	3.70	6.36	5.86

[a]The specified standard errors in the classical calculation (column 3) correspond to the hysteresis of independent forward and backward classical TI calculations of the corresponding systems. The specified standard errors in the QM calculations (columns 4−8) correspond to the standard error of the mean resulting from the accepted HMC runs with QM corrections. The "RMS error" and "max error" correspond to the respective differences between the calculated values and the experimental values.

DOI: 10.1021/acs.jpcb.5b01625
J. Phys. Chem. B 2015, 119, 7030−7040

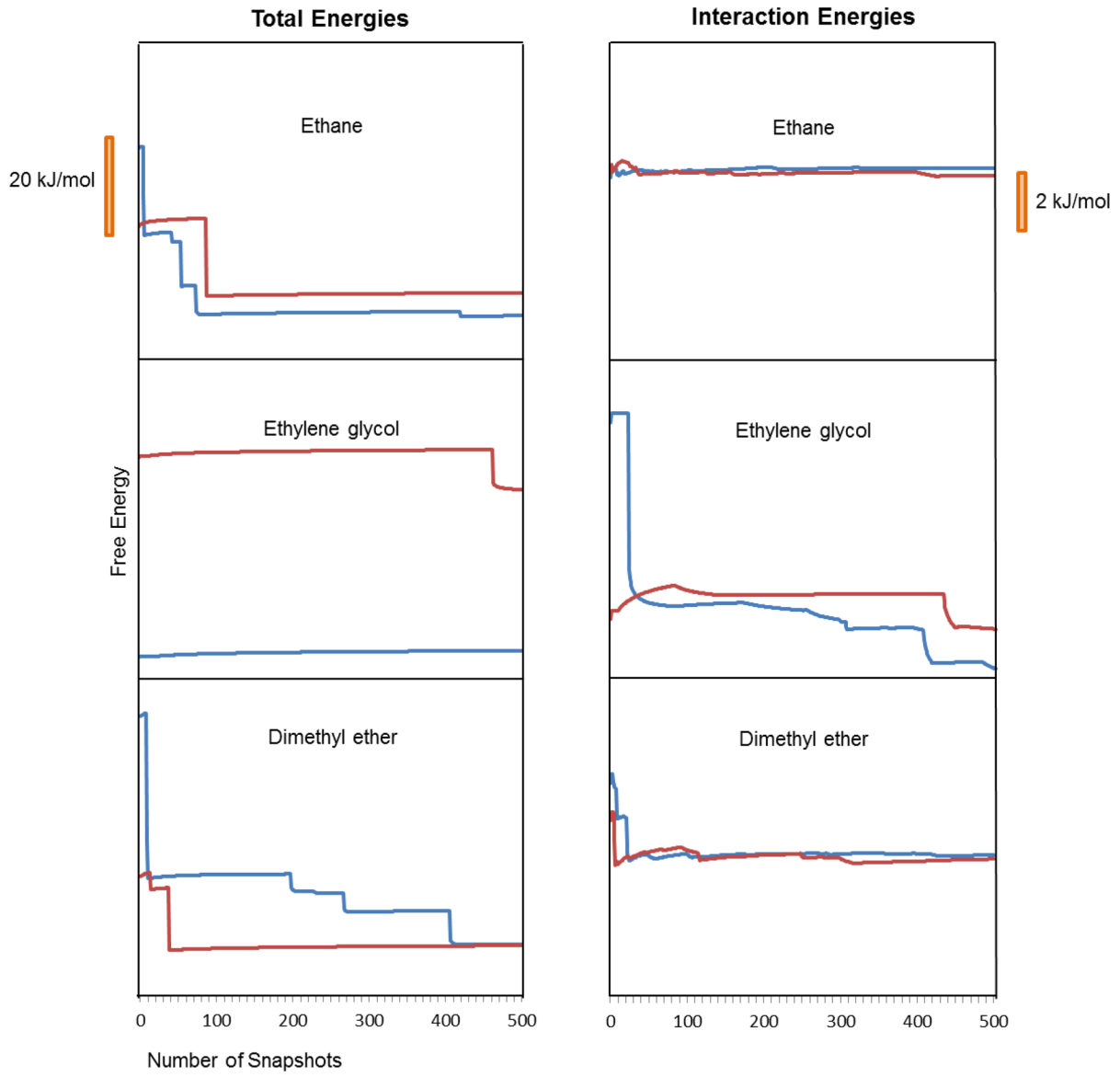

Figure 5. Free energy calculated with the Zwanzig equation as a function of the number of snapshots included when perturbing between the MM +QM^{int}_{Coul} and QM ensembles for ethylene glycol, dimethyl ether, and ethane. The blue line represents run 1, and the red line represents run 2. For clarity, we do not provide an absolute energy scale for the *y* axis, as only the relative energies between each run are important.

intramolecular degrees of freedom, we would expect that the lack of configuration overlap would be caused also by ligand intramolecular terms.

3.5. Interaction Energy Perturbation. Interaction energies are commonly used when applying the Zwanzig equation, as it has often been observed that convergence with total energies is problematic.[23] The smaller magnitudes of interaction energies and the better overlap of the energies lead to better convergence. This can be understood by the notion that the free energy of binding is primarily delivered by the interactions between the ligand and receptor and these are improved by the quantum description.

When using interaction energies to go from QM/MM to full QM (eq 18), a much higher degree of convergence is obtained. This is shown in Figure 5. The largest difference between the two runs is 1.25 kJ/mol (ethylene glycol), well within thermal error (2.5 kJ/mol). These free energies can then be combined with the cycle II free energies from the three-step thermodynamic cycle (Figure 2). The final corrected free energies can be seen in Table 4 under the column labeled "Elstner". In every case, the free energies do not improve. Dimethyl ether and ethylene glycol stay essentially unchanged. Thus, for the more polar ligands, where the main driving force behind the binding energies is electrostatic (e.g., ethylene glycol and dimethyl ether), the correction from QM/MM to QM is small. This is in contrast with the nonpolar ligands, where dispersion is the prominent interaction, where the corrections increase the relative hydration free energies. Overall, the cycle II corrections improve the correlation with the experimental free energies; however, little to no improvement upon cycle II is achieved when applying the final cycle (cycle III) in Figure 2.

We have further examined the method used to calculate the dispersion. In our first attempt to implement cycle III corrections, the Elstner method was used. We have then also examined keeping classical dispersion (as calculated by the force field) and alternatively using Grimme's D2 correction. If the dispersion calculated by the force field is used, it is effectively canceled out within the Zwanzig equation. This is

DOI: 10.1021/acs.jpcb.5b01625
J. Phys. Chem. B 2015, 119, 7030−7040

because the interaction energy at the QM/MM level can be divided into two components, the electrostatics and dispersion.

$$\Delta U^{\text{int}} = \Delta U^{\text{elec}} + \Delta U^{\text{disp}} \quad (19)$$

The electrostatic term came from the QM/MM correction, and the dispersion term comes from the force field. Similarly, the interaction energy at the full QM description can be split into the intrinsic DFT terms (e.g., the Coulomb and exchange) and the empirical dispersion correction. Thus, in the case where we retain the force field dispersion in the QM description, it will cancel out with the dispersion included in QM/MM, as the two are the same.

The use of the force field dispersion significantly improves the free energies calculated for cycle II for every single ligand. Grimme's D2 correction causes the largest errors between the different dispersion approaches, to the point that overall it is a deterioration of accuracy compared to the classical and QM/MM results. We have also tried the more advanced Grimme D3 approach, but the results obtained are essentially the same as those with the D2 method. In order to have a single measure of accuracy for each method, the RMS error for each method was calculated and is shown in Table 4. This was calculated with respect to the experimental values. It is clear that QM/MM and QM with force field dispersion improve upon the classical result, with the QM with force field dispersion producing dramatic improvement with a RMS error of only 2.05 kJ/mol.

Finally, we investigate how our method depends on the choice of exchange-correlation functional. For this purpose, we have tried the LDA and BLYP exchange-correlation functionals in addition to the PBE functional that we have used up to now, for a single mutation (ethanol to ethane), and examined how these changes affect free energies at the QM/MM and full QM descriptions. The results are shown in Table 5.

Table 5. Free Energy Corrections with Different Exchange-Correlation Functionals Applied to the Relative Hydration Free Energy between Ethanol and Ethane[a]

	free energy (kJ/mol)				
functional	MM+QM$^{\text{int}}_{\text{Coul}}$ corrected	Elstner	force field disp	Grimme D2	Grimme D3
PBE	8.71	10.31	7.52	11.07	10.79
LDA	9.46			15.48	
BLYP	8.70	9.07	7.13	9.81	9.24

[a]The experimental relative hydration free energy is 7.66[45] kJ/mol, and the classical free energy is 0.53 kJ/mol.

These results indicate that, within the GGA approximation, switching between PBE and BLYP plays a small role in obtaining accurate corrections. This indicates similar behavior between GGA functionals. For the LDA, the result for the QM/MM description is 0.7 kJ/mol worse than the GGA functionals but really erroneous for the full QM description with an error of about 8 kJ/mol with respect to experiment. We should note that we have not used any dispersion correction for the LDA calculations, as the LDA method intrinsically includes spurious attractive interactions that play a role of dispersion in this case or in any case make the empirical dispersion corrections not applicable to LDA.

Previously (Table 4), the results showed that using the force field dispersion yields the best corrections, and here the same trend can be observed again for PBE and BLYP.

4. CONCLUSIONS

We have presented a "stepping stone" approach for computing QM corrections to MM free energies of binding that aims to overcome the convergence difficulties of similar approaches which are based on a single-step free energy perturbation from the classical to the quantum system. Our approach includes two stages: In the first stage, we gradually mutate the MM system to QM/MM using thermodynamic integration (TI) on intermediate ensembles generated via hybrid Monte Carlo simulations. This stage accommodates most of the change in polarization associated with the MM to QM mutation. As a result, the second stage, which is a single-step QM/MM to full QM mutation, actually converges well. Here, we validated our method on the calculation of hydration free energies for a set of ligands with different polarities. We found that the mutation from MM to QM/MM is a definite improvement in the relative hydration free energies with respect to classical TI results, and the stage 2 correction where we mutate to the full QM ensemble produces further and more substantial improvement, reducing the classical max and RMS errors by a factor of 2. The approach is quite sensitive to the choice of dispersion model but less so in the choice of the GGA exchange correlation functional. We would like to apply this method in future work to compute free energies of binding between proteins and ligands.

AUTHOR INFORMATION

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Chris Sampson would like to thank the BBSRC and Boehringer Ingelheim for an industrial CASE PhD studentship (BBSRC grant reference: BB/I015922/1). The Iridis3 and Iridis4 supercomputers of the University of Southampton were used in this work.

REFERENCES

(1) Beveridge, D.; DiCapua, F. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431−492.
(2) Straatsma, T.; McCammon, J. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407−435.
(3) Gumbart, J.; Roux, B.; Chipot, C. *J. Chem. Theory Comput.* **2013**, *9*, 794−802.
(4) Deng, Y.; Roux, B. *J. Phys. Chem. B* **2009**, *113*, 2234−2246.
(5) Villa, A.; Zangi, R.; Pieffet, G.; Mark, A. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 673−686.
(6) Wang, L.; Berne, B.; Friesner, R. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1937−1942.
(7) Godschalk, F.; Genheden, S.; Söderhjelm, P.; Ryde, U. *Phys. Chem. Chem. Phys.* **2013**, *15*, 7731−7739.
(8) de Ruiter, A.; Oostenbrink, C. *Curr. Opin. Chem. Biol.* **2011**, *15*, 547−552.
(9) Hu, H.; Yang, W. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573−601.
(10) Hayik, S. A.; Liao, N.; Merz, K. M., Jr. *J. Chem. Theory Comput.* **2008**, *4*, 1200−1207.
(11) Gräter, F.; Schwarzl, S.; Dejaegere, A.; Fischer, S.; Smith, J. *J. Phys. Chem. B* **2005**, *109*, 10474−10483.
(12) Hayik, S.; Dunbrack, R.; Merz, K. *J. Chem. Theory Comput.* **2010**, *6*, 3079−3091.
(13) Ihrig, A.; Schiffmann, C.; Sebastiani, D. *J. Chem. Phys.* **2011**, *135*, 214107.
(14) Štrajbl, M.; Hong, G.; Warshel, A. *J. Phys. Chem. B* **2002**, *106*, 13333−13343.

DOI: 10.1021/acs.jpcb.5b01625
J. Phys. Chem. B 2015, 119, 7030−7040

(15) Beierlein, F.; Michel, J.; Essex, J. *J. Phys. Chem. B* **2011**, *115*, 4911−4926.
(16) Zwanzig, R. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(17) Rod, T.; Ryde, U. *J. Chem. Theory Comput.* **2005**, *1*, 1240−1251.
(18) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235−10253.
(19) König, G.; Hudson, P.; Boresch, S.; Woodcock, H. *J. Chem. Theory Comput.* **2014**, *10*, 1406−1419.
(20) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. *J. Chem. Phys.* **2003**, *119*, 3559−3566.
(21) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690−2693.
(22) Jensen, M.; Park, S.; Tajkhorshid, E.; Schulten, K. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 6731−6736.
(23) Fox, S.; Pittock, C.; Tautermann, C.; Fox, T.; Christ, C.; Malcolm, N.; Essex, J.; Skylaris, C.-K. *J. Phys. Chem. B* **2013**, *117*, 9478−9485.
(24) Skylaris, C.-K.; Haynes, P.; Mostofi, A.; Payne, M. *J. Chem. Phys.* **2005**, *122*, 084119.
(25) König, G.; Brooks, B. *Biochim. Biophys. Acta* **2015**, *1850*, 932−943.
(26) Woods, C.; Manby, F.; Mulholland, A. *J. Chem. Phys.* **2008**, *128*, 014109.
(27) Duane, S.; Kennedy, A.; Pendleton, B.; Roweth, D. *Phys. Lett. B* **1987**, *195*, 216−222.
(28) Mehlig, B.; Heermann, D.; Forrest, B. *Phys. Rev. B* **1992**, *45*, 679−685.
(29) Iftimie, R.; Salahub, D.; Wei, D.; Schofield, J. *J. Chem. Phys.* **2000**, *113*, 4852−4862.
(30) Marsaglia, G.; Bray, T. *SIAM Rev.* **1964**, *6*, 260−264.
(31) Hill, Q.; Skylaris, C.-K. *Proc. R. Soc. A* **2009**, *465*, 669−683.
(32) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(33) Wang, J.; Wang, W.; D.C, P. K. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.
(34) Wang, J.; Wolf, R.; Caldwell, J.; Kollman, P.; Case, D. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(35) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926−935.
(36) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(37) Skylaris, C.-K.; Haynes, P.; Mostofi, A.; Payne, M. *Phys. Status Solidi* **2006**, *243*, 973−988.
(38) Wilkinson, K.; Hine, N.; Skylaris, C.-K. *J. Chem. Theory Comput.* **2014**, *10*, 4782−4794.
(39) Skylaris, C.-K.; Mostofi, A.; Haynes, P.; Diéguez, O.; Payne, M. *Phys. Rev. B* **2002**, *66*, 035119.
(40) Mostofi, A.; Haynes, P.; Skylaris, C.-K.; Payne, M. C. *J. Chem. Phys.* **2003**, *119*, 8842−8848.
(41) Haynes, P.; Skylaris, C.-K.; Mostofi, A.; Payne, M. *Chem. Phys. Lett.* **2006**, *422*, 345−349.
(42) Fox, S.; Pittock, C.; Fox, T.; Tautermann, C.; Malcolm, N.; Skylaris, C.-K. *J. Chem. Phys.* **2011**, *135*, 224107.
(43) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149−5155.
(44) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, S. *J. Chem. Phys.* **2010**, *132*, 154104.
(45) Abraham, M.; Whiting, G.; Fuchs, R.; Chambers, E. *J. Chem. Soc., Perkin Trans. 2* **1990**, *1*, 1−300.
(46) Chambers, C.; Hawkins, G.; Cramer, C.; Truhlar, D. *J. Phys. Chem.* **1996**, *100*, 16385−16398.
(47) Cave-Ayland, C.; Skylaris, C.-K.; Essex, J. *J. Phys. Chem. B* **2014**, *119*, 1017−1025.

DOI: 10.1021/acs.jpcb.5b01625
J. Phys. Chem. B 2015, 119, 7030−7040