

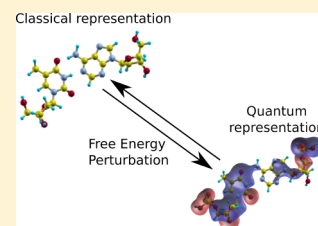
# Direct Validation of the Single Step Classical to Quantum Free Energy Perturbation

Christopher Cave-Ayland, Chris-Kriton Skylaris, and Jonathan W. Essex\*

School of Chemistry, University of Southampton, Highfield, Southampton, Hampshire, SO17 1BJ, United Kingdom

**S** Supporting Information

**ABSTRACT:** The use of the Zwanzig equation in the calculation of single-step perturbations to provide first-principles (*ab initio*) quantum mechanics (QM) correction terms to molecular mechanics (MM) free energy cycles is well established. A rigorous test of the ability to converge such calculations would be very useful in this context. In this work, we perform a direct assessment of the convergence of the MM to QM perturbation, by attempting the reverse QM to MM perturbation. This required the generation of extensive QM molecular dynamics trajectories, using density functional theory (DFT), within the representative biological system of a DNA adenosine–thymidine dimer. Over 100 ps of dynamics with the PBE functional and 6.25 ps with the LDA functional were generated. We demonstrate that calculations with total potential energies are very poorly convergent due to a lack of overlap of phase space distributions between ensembles. While not theoretically rigorous, the use of interaction energies provides far superior convergence, despite the presence of nonclassical charge transfer effects within the DFT trajectories. The source of poor phase space overlap for total energies is diagnosed, the approximate quantification of overlaps suggesting that even for the comparatively simple system considered here convergence of total energy calculations within a reasonable simulation time is unfeasible.



## INTRODUCTION

The accurate and rapid prediction of free energies of binding and hydration for small molecule targets remains a long sought goal in the field of computational chemistry.<sup>1</sup> A range of different techniques have been developed to tackle this problem, the most accurate of which make use of extensive molecular dynamics (MD) or Monte Carlo (MC) sampling and rigorously derived free energy difference estimators.<sup>2–4</sup> Two factors limit the accuracy of these free energy techniques: the realism of the energy model used to describe the potential energy surface of interest and achieving a sufficient degree of sampling of the system to obtain converged ensemble average statistics. Highly realistic, i.e., first-principles quantum mechanics (QM) based energy models are able to accurately model a system's potential energy surface but are prohibitively expensive to undertake sufficient sampling of even moderately sized systems. This practical restriction generally necessitates the use of classically inspired molecular mechanics (MM) force fields. Although computationally far cheaper, the approximate and parametrized nature of MM methods places inherent restrictions on the achievable accuracy of calculations using MM potentials.

The dichotomy between MM and QM approaches has led to the development of hybrid methods that attempt to exploit the accuracy of QM models at a fraction of the computational cost, through judicious combination with MM potentials.<sup>5–9</sup> Perhaps the simplest of these, and the focus of this work, allows the calculation of additional QM correction terms to MM based free energy cycles using a single-step free energy perturbation<sup>10</sup> together with the Zwanzig equation.<sup>2</sup> The one-sided sampling of the Zwanzig equation allows the technique to avoid costly

sampling with the QM Hamiltonian. This advantage is countered however by a more stringent requirement for overlap between perturbation end states than other free energy difference estimators.

The unstable numerical formulation of the Zwanzig equation and its inherent directionality can make it difficult to determine whether the condition of sufficient overlap has been met.<sup>1</sup> In this work, we consider direct assessment of the quality of phase space overlap between MM and QM states through extensive generation of QM MD trajectories to allow calculation of the reverse, QM to MM, perturbation. Comparison of the forward and reverse perturbations between an MM and QM Hamiltonian with the Zwanzig equation allows direct validation of the single-step perturbation procedure used to generate QM corrections.

An adenosine–thymidine DNA base pair is used as a model system, chosen to represent a compromise between biological realism and computational tractability. Previous work has considered the suitability of different MM water models in hybrid calculations;<sup>11</sup> however, the base pair system we consider here provides a far more ambitious and biologically relevant system. The size of the system is sufficient to allow extensive sampling of the QM phase space, while also representing a ubiquitous biological dimer.

Density functional theory<sup>12</sup> (DFT) has arisen as the most common QM method for carrying out MD calculations at the

**Special Issue:** William L. Jorgensen Festschrift

**Received:** June 29, 2014

**Revised:** September 18, 2014

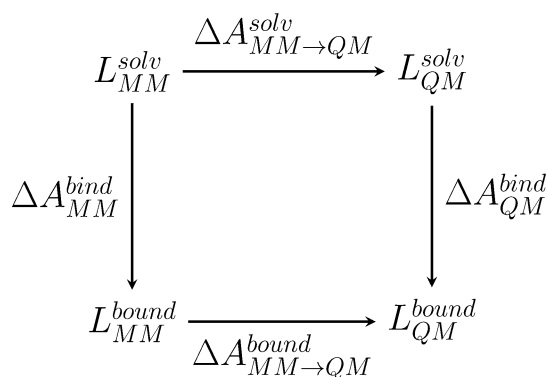
**Published:** September 19, 2014

QM level of theory. Formulations of DFT have been developed that allow scaling to biologically relevant system sizes<sup>9,13,14</sup> (thousands of atoms), dramatically extending the range of systems to which hybrid free energy techniques may be profitably applied.

We generate DFT QM ensembles using MD with the PBE<sup>15</sup> and LDA<sup>12</sup> functionals. The PBE functional has been shown to offer a good compromise between speed and accuracy in describing biological compounds,<sup>16</sup> and is frequently used in this context.<sup>17–21</sup> The LDA functional provides a less realistic description of the system's dynamics but usefully demonstrates the behavior of the single-step perturbation where the MM and QM phase spaces differ more markedly. Classical trajectories are generated using the AMBER ff99SB<sup>22</sup> and GAFF<sup>23</sup> force fields.

## THEORETICAL BACKGROUND

**Single-Step Exponential Averaging.** The QM correction to an MM calculation is given by the thermodynamic cycle in Figure 1. In this case, the free energy of binding at the QM level



**Figure 1.** Free energy cycle for the calculation of QM correction terms to an MM binding free energy difference for the ligand L.

of theory,  $\Delta A_{QM}^{bind}$ , can be obtained from the same calculation at the MM level, combined with the QM correction terms such that  $\Delta A_{QM}^{bind} = \Delta A_{MM}^{bind} - \Delta A_{MM \rightarrow QM}^{solv} + \Delta A_{MM \rightarrow QM}^{bound}$ . The computational cost of inferring  $\Delta A_{QM}^{bind}$  from the cycle must be significantly cheaper than simply calculating this term directly with standard free energy techniques. This is dependent on an efficient method for the calculation of  $\Delta A_{MM \rightarrow QM}^{solv}$  and  $\Delta A_{MM \rightarrow QM}^{bound}$  and is provided through the use of the Zwanzig equation:<sup>2</sup>

$$\Delta A_{0 \rightarrow 1} = -\frac{1}{\beta} \ln \langle \exp[-\beta \Delta U] \rangle_0 \quad (1)$$

The Helmholtz free energy difference between two thermodynamic states 0 and 1 is given by  $\Delta A_{0 \rightarrow 1}$ . Here  $\beta$  has the typical meaning of  $1/k_B T$ , while  $\Delta U = U_1 - U_0$ , i.e., the potential energy difference between the corresponding states, and  $\langle \dots \rangle_0$  represents an ensemble average over state 0. Unlike other commonly used estimators (e.g., TI<sup>4</sup> and BAR<sup>3</sup>), exponential averaging requires sampling of only one end state. In an approach first proposed and employed by Warshel,<sup>10</sup> by choosing the sampled state to be the MM level of theory, it is cheap to generate a series of uncorrelated configurations that can be postprocessed to the QM level:

$$\Delta A_{MM \rightarrow QM} = -\frac{1}{\beta} \ln \langle \exp[-\beta(U_{QM} - U_{MM})] \rangle_{MM} \quad (2)$$

Here  $\Delta A_{MM \rightarrow QM}$  is the free energy difference between the MM and QM descriptions of the same chemical state, the corresponding potential energies denoted by  $U_{QM}$  and  $U_{MM}$ . While free energy calculations are typically broken down into a series of steps using a lambda coupling approach, sampling any intermediate lambda state for an MM to QM perturbation is as prohibitively expensive as sampling under the full QM Hamiltonian.

The Zwanzig equation is notoriously poorly convergent. Calculations using this estimator therefore require a significant degree of phase space overlap between states to converge appropriately.<sup>1</sup> Furthermore, it can be difficult to determine when this criterion has been met—there may be rare, as yet unsampled configurations that will heavily influence the calculated free energy difference.

The drawbacks of the Zwanzig equation should inspire considerable caution, and to our knowledge, it has yet to be rigorously demonstrated that in general the overlap of QM and MM free energy surfaces is sufficient to allow its use. Previous work from this group has developed an alternative approach based around charge perturbation to test for convergence of hybrid MM and QM calculations.<sup>6</sup> This provides a necessary but not sufficient condition for convergence. We directly address the convergence of single step perturbations by considering the calculation of the reverse QM to MM process. As free energy is a state property, the free energy difference between the MM and QM states is invariant based on the direction of the calculation. This provides a rigorous test for convergence based on the condition

$$\Delta A_{MM \rightarrow QM} + \Delta A_{QM \rightarrow MM} = 0 \quad (3)$$

In addition to the previously defined  $\Delta A_{MM \rightarrow QM}$ , the reverse perturbation from the QM to the MM state is denoted by  $\Delta A_{QM \rightarrow MM}$ . Evaluation of both terms in eq 3 applied to a model system under different Hamiltonians therefore provides a direct assessment of the feasibility of and degree of sampling required in converging hybrid free energy calculations. Throughout this work, the deviation from zero of eq 3 will be referred to as the *discrepancy* of a perturbation.

**Interaction Free Energy Differences.** It is common practice when employing hybrid free energy techniques to make use of interaction energies in the place of total energies within the free energy difference estimator.<sup>6,7,9–11,24–29</sup> The interaction energy of a system,  $U_{AB}^{inter}$ , is given by

$$U_{AB}^{inter} = U_{AB} - U_A - U_B \quad (4)$$

where A and B denote two different components of the system, the interaction energy of the two is given by the energy of the complex,  $U_{AB}$ , minus the energy of the two components in isolation,  $U_A$  and  $U_B$ . In the case of a typical MM model, the interaction energy can simply be derived by summing the appropriate terms of the force field, while for QM models additional calculations are required to account for the polarization effects within the complex. Interaction energies are then simply substituted in the place of total energies within the estimator. In the case of the Zwanzig equation

$$\Delta A_{0 \rightarrow 1} = -\frac{1}{\beta} \ln \langle \exp[-\beta \Delta U^{inter}] \rangle_0 \quad (5)$$

This substitution is not without theoretical difficulties, as the derivation of the Zwanzig equation is carried out using total energies. As such, the consequences of this approximation are unclear. As defined above, the interaction energy of a system includes the energy of polarization and hence free energy calculations using interaction energies are still able to capture these effects.

**Phase Space Overlap.** The degree of phase space overlap between thermodynamic states was assessed directly using the following metric based on Bennett's acceptance ratio (BAR).<sup>3</sup>

$$\Delta A_{0 \rightarrow 1} = \frac{1}{\beta} \ln \frac{\langle f(U_0 - U_1 + C) \rangle_1}{\langle f(U_1 - U_0 - C) \rangle_0} + C - \frac{1}{\beta} \ln \frac{N_1}{N_0} \quad (6)$$

$$C = \Delta A_{0 \rightarrow 1} + \frac{1}{\beta} \ln \frac{N_1}{N_0} \quad (7)$$

Here  $f(x)$  is the Fermi function  $f(x) = (1 + \exp[\beta x])^{-1}$ , while  $N_0$  and  $N_1$  are the number of snapshots used to calculate the respective ensemble averages. Equations 6 and 7 are solved self-consistently until a converged estimate for  $\Delta A_{0 \rightarrow 1}$  is obtained. The value of  $C$  describes an arbitrary shift in the relative height of the two potential energy surfaces under consideration, the self-consistent procedure giving the optimum value of  $C$  than minimizes the statistical uncertainty of  $\Delta A_{0 \rightarrow 1}$ . At this value of  $C$ , denoted here as  $C^{\text{opt}}$ , the condition  $\langle f(U_0 - U_1 + C^{\text{opt}}) \rangle_1 = \langle f(U_1 - U_0 - C^{\text{opt}}) \rangle_0$  is met.

Bennett notes that the value to which the ensemble averages converge, we shall refer to this as  $O_{\text{BAR}}$ , provides information about the sufficiency of sampling within a calculation, and is given by

$$O_{\text{BAR}} = \langle f(U_0 - U_1 + C^{\text{opt}}) \rangle_1 = \langle f(U_1 - U_0 - C^{\text{opt}}) \rangle_0 \quad (8)$$

$O_{\text{BAR}}$  being small indicates that insufficient sampling of important regions of phase space has occurred, whereas values approaching 1 indicate sufficiency of sampling. Hence  $O_{\text{BAR}}$  may be used to assess the relative overlaps in phase space of different potential energy surfaces by comparing values between calculations with similar levels of sampling. Consider two different free energy calculations of similar length, where one gives a large value of  $O_{\text{BAR}}$  but for which the other is small. It may be reasonably concluded that as each calculation has been sampled equivalently the perturbation with the larger value for  $O_{\text{BAR}}$  displays better overlap in phase space.

Additionally, Bennett provides an expression for direct calculation of phase space overlap in the form of the following integral:

$$O = 2 \int \frac{P_0(\mathbf{x})P_1(\mathbf{x})}{P_0(\mathbf{x}) + P_1(\mathbf{x})} d\mathbf{x} \quad (9)$$

where  $P_0(\mathbf{x})$  and  $P_1(\mathbf{x})$  give the probability of a configuration  $\mathbf{x}$  under different ensembles. Owing to the unfeasibility of evaluating integrals of more than a few dimensions, we make use of this expression in only a few single dimensional cases to estimate phase space overlap of particular degrees of freedom.

## METHODS AND CALCULATION SETUP

**QM Calculations.** All QM calculations were carried out using the plane-wave DFT package CASTEP 5.5.<sup>30</sup> Calculations using the LDA<sup>12</sup> functional were carried out with a kinetic energy cutoff of 900 eV with norm-conserving pseudopotentials.<sup>31</sup> PBE<sup>15</sup> calculations used a kinetic energy

cutoff of 500 eV and ultrasoft pseudopotentials automatically generated by CASTEP. Kinetic energy cutoffs in each case were tested and chosen on the basis of the requirement of converged energies. Electronic energies were converged to a tolerance of  $10^{-5}$  eV per atom between SCF cycles, using a maximum g-vector of  $0.1 \text{ \AA}^{-1}$  for charge mixing and a grid spacing factor of 2.0 relative to the diameter of the cutoff sphere. A cubic periodic box with sides of 20 Å was used for LDA calculations and 25 Å for PBE; both box sizes are more than sufficient to accommodate the A–T dimer. Long range electrostatics were treated through Ewald summation.

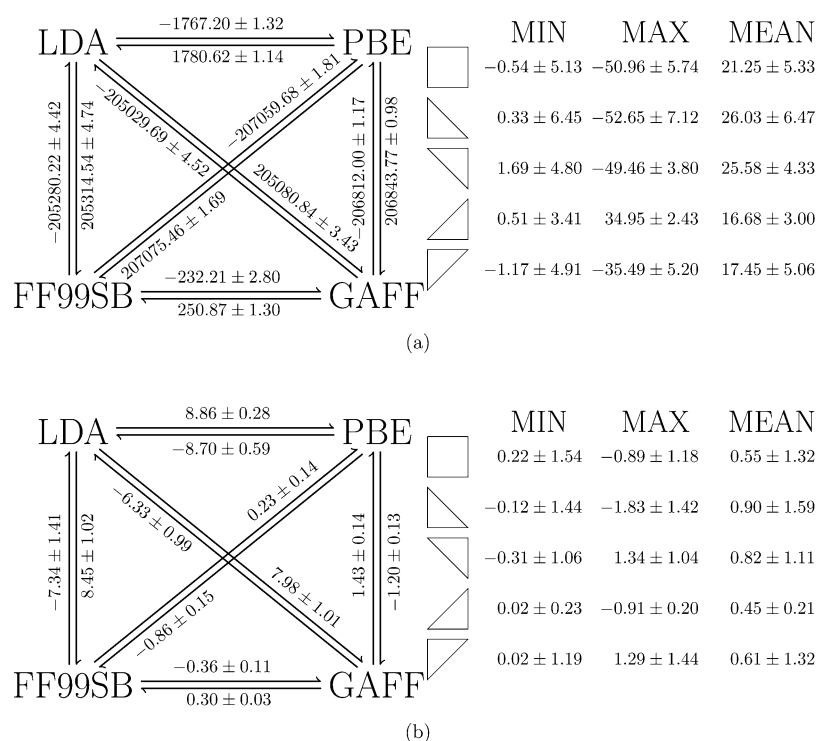
**MM Calculations.** All MM calculations were carried out using the AMBER 12 software suite.<sup>32,33</sup> Calculations were carried out using both the GAFF<sup>23</sup> and ff99SB<sup>22</sup> force fields. Partial charges for use with the GAFF force field were produced with ANTECHAMBER using the AM1-BCC charge method.<sup>34,35</sup> A cutoff of 8 Å was used in the calculation of nonbonded interactions, and the particle mesh Ewald (PME) method was used for long-range electrostatics. The PME was validated against the conventional Ewald approach for electrostatics to confirm the equivalent treatment between the MM and QM Hamiltonians (see the Supporting Information). A cubic periodic box with sides of 20 Å was used for ff99SB calculations and 25 Å for GAFF.

**Molecular Dynamics.** The same MD protocol was used for both the MM and QM systems. Initial structures for production MD were generated by the NAB module of AMBER, and subsequently minimized for 50 iterations with the appropriate potential energy function. Bases were modeled with the associated deoxyribose component but without phosphate present. For each Hamiltonian, five independent repeats with the same starting configuration were run. All MD calculations were carried out with a time step of 0.25 fs, as determined by the requirement for constant energy dynamics under the NVE ensemble. Production MD runs were carried out in the NVT ensemble with periodic boundary conditions. Temperature control was achieved using the Langevin thermostat with a collision constant of  $0.1 \text{ ps}^{-1}$  to regulate the system at 300 K.

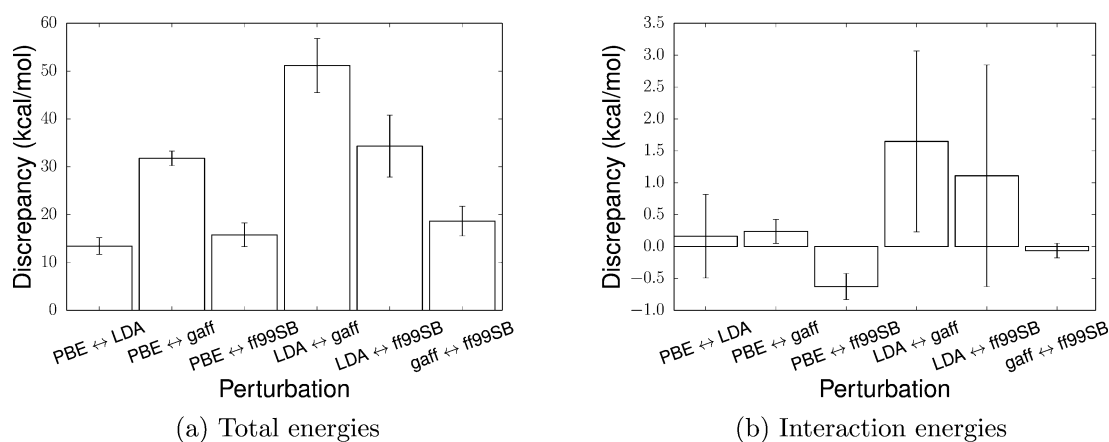
The only differences between MD calculations for the MM and QM systems, besides the choice of Hamiltonian, lies in the different algorithms used by CASTEP and AMBER. AMBER employs the leapfrog integrator to solve equations of motion, while minimizations employed the conjugate gradient algorithm. Born–Oppenheimer *ab initio* MD calculations in CASTEP employed the velocity Verlet algorithm, while minimizations were based on the Broyden–Fletcher–Goldfarb–Shannon (BFGS) algorithm.<sup>36</sup>

Generation of QM and MM trajectories was carried out simultaneously and was continued until the discrepancy of all perturbations with interaction energies was close to zero. This criteria produced a total trajectory length of 6.25 ps with the LDA functional and 100.0 ps with the PBE functional. In each case, this total simulation time was split between five independent repeats of equal length. The only exception to this is the ff99SB trajectories that do not match the full length of the PBE trajectories but are 25 ps in total.

**Potential of Mean Force Calculations.** Potential of mean force calculations were carried out using MD with linear constraints with CASTEP.<sup>37</sup> An additional 25 short (1500 time steps) MD runs were carried out with linear constraints placed on the N–H–N hydrogen bond between thymidine and adenosine. The N–H and H–N bonds were considered as separate degrees of freedom constrained at 0.2 Å intervals, from



**Figure 2.** Free energy cycles constructed between all Hamiltonians using (a) total energies and (b) interaction energies. A single standard error for each perturbation is shown, derived from the standard deviation of the five repeats of each calculation. On the right of each diagram, the minimum, maximum, and mean closures of the illustrated cycle are shown. Standard errors for closures are calculated by summing the variance of each leg of the cycle involved. Standard errors for mean closures were calculated by taking the average variance of all possible leg permutations for each cycle.



**Figure 3.** Discrepancies for forward and reverse perturbations within the free cycles using (a) total energies and (b) interaction energies. One standard error is shown for all results, calculated by summing the variance of the forward and reverse calculation.

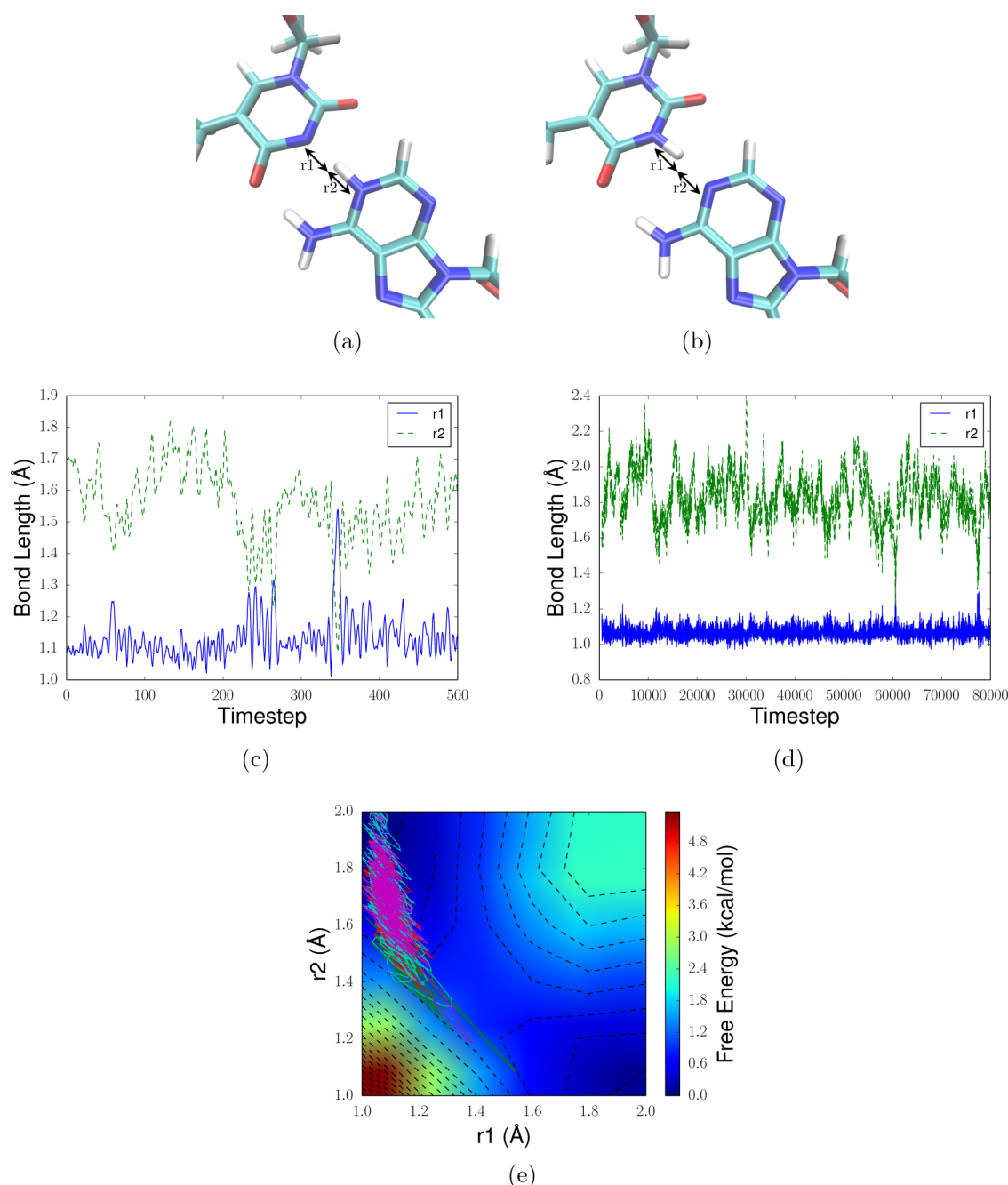
1.0 to 1.8 Å. This gives a 5 by 5 grid of points, corresponding to the 25 runs. Constraints were enforced using the RATTLE<sup>38</sup> algorithm. The mean force required to maintain each constraint is equal to the gradient of the free energy surface at that point. The surface itself is then generated through use of the Euler method,<sup>39</sup> taking the lowest point of the PMF to be zero.

## RESULTS AND DISCUSSION

**Single-Step FEP.** Results of all perturbations between the four considered Hamiltonians are given in Figure 2. Although our stated aim was to examine perturbations between MM and QM states, it was considered trivial additional work to complete the calculations for all possible perturbations. Completion of the larger cycle allows for a more rigorous test of convergence

through the computation of cycle closures. Unfortunately, closures are nontrivial to calculate, as each leg has two separately calculated free energy differences associated with it. Different forward and reverse calculations can be used in any permutation to provide a value for the cycle closure. We compromise by calculating all possible permutations for each cycle and reporting the minimum, maximum, and mean unsigned closures. It is immediately apparent from these results that interaction energies provide much tighter cycle closures than using total energies. Although the reported minimum closures using total energies are close to zero, the large associated standard errors suggest this is simply spurious, through a fortunate combination of different components of the cycle. The mean and maximum closures are exceedingly poor





**Figure 4.** (a and b) Example configurations from an LDA MD run, with the proton exchanged (a) and not exchanged (b). (c and d) Time series of  $r_1$  and  $r_2$  from an example LDA (c) and PBE (d) MD run. (e) Free energy surface of proton exchange between bases using the LDA functional. The solid lines indicate the paths taken by the five LDA MD trajectories. Dashed contour lines are plotted every 0.25 kcal/mol.

however and suggest the unsuitability of total energies in hybrid free energy work. Although a recent paper has presented results that give successful convergence with total energy calculations, the general applicability of this approach has yet to be demonstrated in a system as complex as that considered here.<sup>8</sup>

The convergence of each leg of the cycle can be assessed by calculating the discrepancy between the forward and reverse perturbation, as given by eq 3. This is shown in Figure 3. For total energies, no particular leg in the cycle can be highlighted as responsible for the poor convergence; even the best converged leg (the PBE to LDA perturbation) has a discrepancy of greater than 10 kcal/mol. The use of interaction energies however is much more compelling. All perturbations

fall close to or less than one standard error from zero, with the exception of the PBE to ff99SB calculations ( $p$ -value < 0.05 from an unpaired Student's  $t$  test that the free energy differences in either direction are drawn from different distributions).

The magnitude of the free energy differences is considerable when using total energies. Interpretation of these values should be taken with care, as all energies calculated are given with respect to an arbitrary reference value, determined by the Hamiltonian. The difference in this reference value between Hamiltonians gives very large apparent free energy differences. The use of the free energy cycles such as in Figure 1 accounts for this reference state effect and gives meaningful relative free

energy changes. Interpretation of the values associated with individual legs of the cycle should be carried out with care however.

It might be argued that the convergence of calculations using total energies fails simply due to the large differences in the size of the energy values associated with each Hamiltonian. To test for the possibility of numerical instability caused by differences in reference state, arbitrary constants were used to adjust energy values within individual perturbations. This allows for the exponential terms in Figure 2 to be scaled to numerically tractable regions; the unadjusted free energy difference can then be recovered by removing the arbitrary constant used. In practice, this procedure was found to have no effect on the discrepancy of each perturbation. Moreover, it can be shown analytically that the discrepancy is invariant with respect to the difference in reference state between Hamiltonians (see the Supporting Information). As long as care is taken to avoid numerical overflows in the exponential terms, the difference in scale of the energy values has no effect on the convergence properties of a calculation.

These results indicate that, in practical terms, the use of single step perturbation techniques is restricted to interaction energies. In addition to the significantly superior convergence properties of interaction energies, they provide a more intuitive interpretation for the resulting free energy differences, as differences in the strength of interaction under different Hamiltonians. For interaction energies, all Hamiltonians share a naturally defined common reference state, namely, the two bases at infinite separation. In practice, the use of interaction energies is commonplace with hybrid MM and QM work.<sup>6,7,9–11,24–29</sup> Despite this prevalence, however, it is our opinion that the use of interaction energies is not formally correct in the context of free energy calculations based on the Zwanzig equation which is derived for total energies. A rigorous theoretical and practical examination of the consequences of using interaction energies will be presented in upcoming work. In practice, however, the poor convergence of total energy calculations leaves little choice but to use interaction energies.

The failure of calculations using total energies is suggestive of poor overlap between the potential energy surfaces of the different Hamiltonians. That only total energies are affected suggests the problem pertains to the intramolecular degrees of freedom of the system. This is considered in more detail in a later section.

**QM MD Trajectories.** Within the QM trajectories, some examples of proton exchange were observed between the N–H of the thymidine and the N hydrogen bonding partner of adenosine (see Figure 4). Marked exchange events were observed within two of the five trajectories with the LDA functional; this is particularly significant given their short duration. In contrast, the PBE functional demonstrated comparatively little exchange, only two events occurring within one the five repeats of considerably greater length. Characterization of the free energy barrier of proton exchange under the LDA functional was carried out through potential of mean force of constraint (PMFC) calculations, using CASTEP. This reveals a free energy barrier of around 1.0 kcal/mol, well within the range expected to be crossed due to thermal fluctuations at 300 K. This value is perhaps underestimated due to the coarse resolution of the PMF and the short, constrained trajectories used to generate it. The key features of the landscape appear to be recreated, however, and transitions between the minima occur across the saddle point. The observation of hydrogen

exchange within this system may also be attributed to the propensity of DFT functionals to underestimate proton exchange barriers.<sup>40</sup>

The comparative rarity of proton exchange events under the more accurate PBE functional suggests that exchange is due to the shortcomings of the LDA functional, leading to unphysically low barriers within the MD runs. Production of an LDA ensemble is still of considerable value, as it is noted that a converged free energy difference can still be calculated even where the QM Hamiltonian includes nonclassical effects, such as charge transfer or polarization. Owing to the formulation of the Zwanzig equation, configurations with very high energies in the classical Hamiltonian (such as a highly stretched covalent bond in the case of the proton exchange) are negligibly likely to occur under classical dynamics and hence do not contribute to the free energy difference. Conversely, while sampling under the QM Hamiltonian, configurations stabilized by nonclassical effects have large negative values of  $\Delta U$  and hence small contributions to the overall free energy difference.

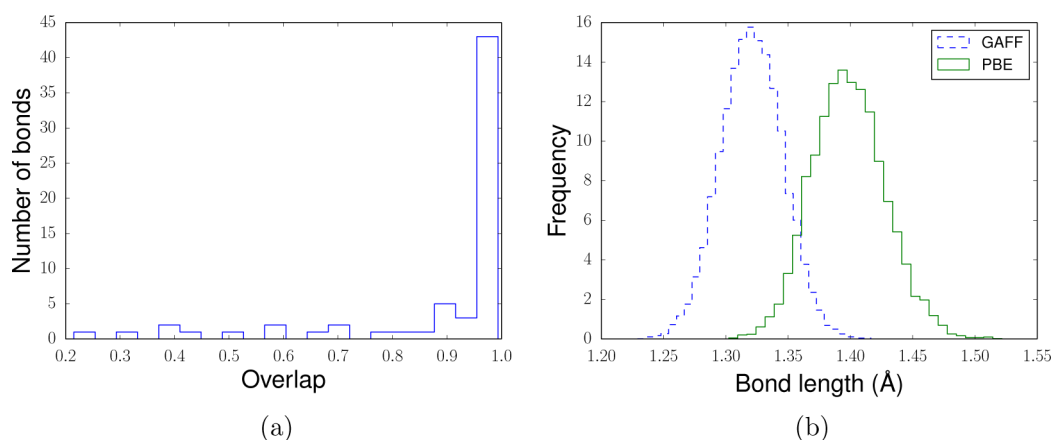
**Phase Space Overlap.** The failure of calculations to converge with the use of total energies is indicative of a violation of the requirement for sufficient phase space overlap of not only the MM and QM potential energy surfaces but of all the energy models. That this problem can be ameliorated with the use of interaction energies suggests the practical reason for the widespread use of this approximation. Normal modes show very good agreement between all Hamiltonians used in this work, suggesting that normal-mode analysis is insufficient to assess phase space similarity in this case (see the Supporting Information).

To examine the extent to which using interaction energies improves phase space overlap, the value of  $O_{\text{BAR}}$  was calculated for all perturbations using total and interaction energies (Table 1). Although values of  $O_{\text{BAR}}$  cannot profitably be compared

**Table 1.**  $O_{\text{BAR}}$  Values for Each Perturbation Using Total and Interaction Energies, Calculated as Described in the Theoretical Background Section

perturbation	total energies	interaction energies
GAFF $\leftrightarrow$ ff99SB	$3.71 \times 10^{-3} \pm 3.72 \times 10^{-3}$	$0.98 \pm 0.00$
LDA $\leftrightarrow$ ff99SB	$4.38 \times 10^{-10} \pm 8.75 \times 10^{-10}$	$0.12 \pm 0.04$
LDA $\leftrightarrow$ GAFF	$1.46 \times 10^{-16} \pm 2.91 \times 10^{-16}$	$0.08 \pm 0.04$
PBE $\leftrightarrow$ ff99SB	$5.46 \times 10^{-5} \pm 1.20 \times 10^{-4}$	$0.56 \pm 0.02$
PBE $\leftrightarrow$ GAFF	$9.77 \times 10^{-12} \pm 1.38 \times 10^{-11}$	$0.40 \pm 0.03$
PBE $\leftrightarrow$ LDA	$2.19 \times 10^{-5} \pm 1.76 \times 10^{-5}$	$0.56 \pm 0.14$

between perturbations due to differing simulation lengths, values for total and interaction energies within perturbations can be compared directly, as they are produced from the same data. The use of interaction energies provides between 5 and 16 orders of magnitude improvement in the value  $O_{\text{BAR}}$ . The smaller overlap values for interaction energies between the LDA functional and classical potentials can be rationalized in terms of the proton-exchange events seen in the LDA trajectories. It is comforting to note that the calculated overlap is superior between the PBE functional and classical potentials than to the LDA functional. Perhaps unsurprisingly, the specialized parameters of the ff99SB force field are noted to offer enhanced overlap with the PBE functional compared to the GAFF force field. Regardless, the values of  $O_{\text{BAR}}$  presented for perturbations involving GAFF are still more than sufficient to suggest the feasibility of the single-step perturbation.



**Figure 5.** (a) Distribution of overlaps between GAFF and PBE ensembles for covalent bonds in the base pair calculated using eq 9. (b) Distribution of the C4–O4 bond of thymidine under different Hamiltonians.

The striking improvement in phase space overlap provided by interaction energies suggests that the poor total energy results are likely due to the failures in the overlap of intramolecular degrees of freedom. Using interaction energies reduces the number of degrees of freedom that are considered within the perturbation to exclude intramolecular terms. Additionally, it is noted generally that intramolecular potentials tend to be less “soft” than their intermolecular counterparts. This suggests that in general it is easier to satisfy the required phase space overlap for intermolecular interactions that have broader probability distributions.

To pinpoint the particular intramolecular degrees of freedom that give rise to poor total energy overlaps, a simple analysis restricted to the systems’ bond lengths was used. From the trajectory data, distributions for all bond lengths under the PBE and GAFF Hamiltonians were generated, as these are the longest and hence best sampled trajectories. Overlaps between the distributions of corresponding bonds under different Hamiltonians were then calculated using eq 9. The 64 covalent bonds in the base pair give rise to a distribution of overlaps, as shown in Figure 5a. The majority of bonds display excellent overlap between the MM and QM ensembles, but a number demonstrate considerably reduced overlap caused by an offset in equilibrium lengths. The worst example of this is given in Figure 5b, showing the C4–O4 bond within thymidine (using the Amber force field atom naming conventions<sup>22</sup>).

Each bonded degree of freedom can be approximated as varying independently with respect to the other bonds of the system (see the Supporting Information for correlation analysis of MD trajectories). An estimate of the combined overlap of the PBE and GAFF Hamiltonians can therefore be obtained by taking the product of the overlaps for each individual bond. This overlap estimate is limited to a subregion of the configuration space of the system as defined by those covalently bonded degrees of freedom and gives a value of  $2.132 \times 10^{-5}$ . This represents a generous upper bound on the overlap of the two states, as the inclusion of additional degrees of freedom can only serve to lower the combined overlap of the system. Although the majority of bonds within the system present an overlap of greater than 0.95, the comparatively small number with poor overlap values can combine to give a globally poor overlap between states. This estimate of the overlap falls short of that required for the convergence of calculations using BAR.<sup>41</sup> As a less efficient estimator, the Zwanzig equation requires even better phase space overlap between states. The

values of  $O_{\text{BAR}}$  for the different perturbations presented in Table 1 support the use of the Zwanzig equation, as they suggest significant overlap is achieved between intermolecular degrees of freedom.

## CONCLUSIONS

The data presented in this work constitute a direct validation of the MM to QM single step free energy perturbation procedure, through completion of the reverse QM to MM perturbation. This required the generation of extensive *ab initio* MD trajectories within a model biological system. The A–T DNA dimer chosen for these calculations represents a compromise between biological complexity and expense of calculations. In total, over 100 ps of *ab initio* MD was generated using plane-wave DFT.

Importantly, the practical restriction that perturbations must be carried out with interaction energies instead of total potential energies is established. Discrepancies between forward and reverse perturbations are shown to be on the order of tens of kcal/mol, for total energies, but nearing zero for interaction energies. Although single step perturbation techniques have been used for some time, the requirement to use interaction energies is often glossed over or not explicitly stated.

The failure of total energy calculations with the Zwanzig equation is explained in terms of poor phase space overlap between MM and QM Hamiltonians. Marked differences between the phase space distributions of intramolecular degrees of freedom are highlighted as problematic. Although limited to only the covalently bonded degrees of freedom, our analysis gives very low upper-bound estimates for total energy phase space overlap. This analysis also suggests caution in hybrid free energy work around the common practice of enforcing bond length constraints. Constraints may improve overlap between MM and QM ensembles, by removing problematic degrees of freedom from being sampled, but run the risk of constraining ensembles outside their global minimum, distorting calculated free energy differences. The extent to which this problem may be avoided through the use of interaction energies is unclear. König et al. have examined the effect of bond length constraints in a simple hybrid free energy perturbation of ethane to methanol.<sup>8</sup>

Interaction energy calculations are demonstrated to exhibit markedly better overlap between ensembles, and improved convergence of single step free energy calculations. The

presence of nonclassical proton exchange interactions between the bases does not prevent stable convergence of the calculated free energy differences.

In addition to the single step methodology considered in this work, there have been other notable suggestions for hybrid free energy work based around more elaborate sampling or reweighting techniques.<sup>7,8</sup> We expect the generated QM ensembles from this work to provide a valuable data set for the analysis of other methodologies.

## ■ ASSOCIATED CONTENT

### ● Supporting Information

Analytical proof that discrepancy of a perturbation is invariant to the use of different reference states. Correlation analysis for covalently bonded degrees of freedom from GAFF and PBE trajectories. Comparison of interaction energies calculated for configurations with PME and direct Ewald sums to demonstrate consistency of the electrostatics between AMBER and CASTEP. Comparison of initial minimized configurations used for MD. Normal mode analysis of initial MD configurations. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: J.W.Essex@soton.ac.uk.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by an EPSRC Doctoral Training Centre grant (EP/G03690X/1). The authors would like to gratefully thank the Institute for Complex Systems Simulation Doctoral Training Centre. Calculations in this work made use of the Iridis3 and Iridis4 Supercomputers from the University of Southampton. Additional calculations were performed on the UK National Supercomputing Service, using both HECToR and ARCHER.

## ■ REFERENCES

- (1) Michel, J.; Essex, J. W. Prediction of protein–ligand binding affinity by free energy simulations: assumptions, pitfalls and expectations. *J. Comput.-Aided Mol. Des.* **2010**, *24* (8), 639–658.
- (2) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22* (8), 1420–1426.
- (3) Bennett, C. H. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (4) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3* (5), 300–313.
- (5) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103* (2), 227–249.
- (6) Beierlein, F. R.; Michel, J.; Essex, J. W. A Simple QM/MM Approach for Capturing Polarization Effects in Protein–Ligand Binding Free Energy Calculations. *J. Phys. Chem. B* **2011**, *115* (17), 4911–4926.
- (7) Woods, C. J.; Manby, F. R.; Mulholland, A. J. An Efficient Method for the Calculation of Quantum Mechanics/Molecular Mechanics Free Energies. *J. Chem. Phys.* **2008**, *128* (1), 014109.
- (8) König, G.; Hudson, P. S.; Boresch, S.; Woodcock, H. L. Multiscale Free Energy Simulations: An Efficient Method for Connecting Classical MD Simulations to QM or QM/MM Free Energies Using Non-Boltzmann Bennett Reweighting Schemes. *J. Chem. Theory Comput.* **2014**, *10* (4), 1406–1419.
- (9) Fox, S. J.; Pittock, C.; Tautermann, C. S.; Fox, T.; Christ, C.; Malcolm, N. O. J.; Essex, J. W.; Skylaris, C. K. Free Energies of Binding from Large-Scale First-Principles Quantum Mechanical Calculations: Application to Ligand Hydration Energies. *J. Phys. Chem. B* **2013**, *117* (32), 9478–9485.
- (10) Vaidehi, N.; Wesolowski, T. A.; Warshel, A. Quantum-Mechanical Calculations of Solvation Free Energies. A Combined *Ab Initio* Pseudopotential Free-Energy Perturbation Approach. *J. Chem. Phys.* **1992**, *97* (6), 4264–4271.
- (11) Shaw, K. E.; Woods, C. J.; Mulholland, A. J. Compatibility of Quantum Chemical Methods and Empirical (MM) Water Models in Quantum Mechanics/Molecular Mechanics Liquid Water Simulations. *J. Phys. Chem. Lett.* **2010**, *1* (1), 219–223.
- (12) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, 1133–1138.
- (13) Bowler, D. R.; Miyazaki, T.; Gillan, M. J. Recent Progress in Linear Scaling *Ab Initio* Electronic Structure Techniques. *J. Phys.: Condens. Matter* **2002**, *14* (11), 2781–2798.
- (14) Skylaris, C. K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. Introducing ONETEP: Linear-Scaling Density Functional Simulations on Parallel Computers. *J. Chem. Phys.* **2005**, *122* (8), 084119.
- (15) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (16) Kaschner, R.; Hohl, D. Density Functional Theory and Biomolecules: A Study of Glycine, Alanine, and Their Oligopeptides. *J. Phys. Chem. A* **1998**, *102* (26), 5111–5116.
- (17) Fox, S.; Wallnoefer, H.; Fox, T.; Tautermann, C.; Skylaris, C. First Principles-Based Calculations of Free Energy of Binding: Application to Ligand Binding in a Self-Assembling Superstructure. *J. Chem. Theory Comput.* **2011**, *7*, 1102–1108.
- (18) Robinson, M.; Haynes, P. D. Dynamical Effects in *Ab Initio* NMR Calculations: Classical Force Fields Fitted to Quantum Forces. *J. Chem. Phys.* **2010**, *133* (8), 084109.
- (19) Fonseca Guerra, C.; van der Wijst, T.; Poater, J.; Swart, M.; Bickelhaupt, F. Adenine Versus Guanine Quartets in Aqueous Solution: Dispersion-Corrected DFT Study on the Differences in  $\pi$ -Stacking and Hydrogen-Bonding Behavior. *Theor. Chem. Acc.* **2010**, *125* (3–6), 245–252.
- (20) Schwegler, E.; Galli, G.; Gygi, F. Conformational dynamics of the dimethyl phosphate anion in solution. *Chem. Phys. Lett.* **2001**, *342* (3–4), 434–440.
- (21) Fox, S. J.; Pittock, C.; Fox, T.; Tautermann, C. S.; Malcolm, N.; Skylaris, C. K. Electrostatic Embedding in Large-Scale First Principles Quantum Mechanical Calculations on Biomolecules. *J. Chem. Phys.* **2011**, *135* (22), 224107.
- (22) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21* (12), 1049–1074.
- (23) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.
- (24) Liu, W.; Sakane, S.; Wood, R. H.; Doren, D. J. The Hydration Free Energy of Aqueous Na<sup>+</sup> and Cl<sup>−</sup> at High Temperatures Predicted by *ab Initio*/Classical Free Energy Perturbation: 973 K with 0.535 g/cm<sup>3</sup> and 573 K with 0.725 g/cm<sup>3</sup>. *J. Phys. Chem. A* **2002**, *106* (7), 1409–1418.
- (25) Rod, T. H.; Rydberg, P.; Ryde, U. Implicit Versus Explicit Solvent in Free Energy Calculations of Enzyme Catalysis: Methyl Transfer Catalyzed by Catechol O-Methyltransferase. *J. Chem. Phys.* **2006**, *124* (17), 174503.
- (26) Heimdahl, J.; Rydberg, P.; Ryde, U. Protonation of the Proximal Histidine Ligand in Heme Peroxidases. *J. Phys. Chem. B* **2008**, *112* (8), 2501–2510.
- (27) Wesolowski, T.; Warshel, A. *Ab Initio* Free Energy Perturbation Calculations of Solvation Free Energy Using the Frozen Density Functional Approach. *J. Phys. Chem.* **1994**, *98* (20), 5183–5187.



- (28) Wood, R. H.; Yezdimer, E. M.; Sakane, S.; Barriocanal, J. A.; Doren, D. J. Free Energies of Solvation with Quantum Mechanical Interaction Energies from Classical Mechanical Simulations. *J. Chem. Phys.* **1999**, *110* (3), 1329–1337.
- (29) Genheden, S.; Cabedo Martinez, A.; Criddle, M.; 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* (3), 187–200.
- (30) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I. J.; Refson, V. K. First Principles Methods Using CASTEP. *Z. Kristallogr.* **2005**, *220*, 567–570.
- (31) Segall, M. D.; Lindan, P. J. D.; Probert, M. J.; Pickard, C. J.; Hasnip, P. J.; Clark, S. J.; Payne, M. C. First-Principles simulation: Ideas, Illustrations and the CASTEP code. *J. Phys.: Condens. Matter* **2002**, *14* (11), 2717–2744.
- (32) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*; University of California: San Francisco, CA, 2012.
- (33) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688.
- (34) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. *J. Comput. Chem.* **2000**, *21* (2), 132–146.
- (35) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23* (16), 1623–1641.
- (36) Pfrommer, B. G.; Côté, M.; Louie, S. G.; Cohen, M. L. Relaxation of Crystals with the Quasi-Newton Method. *J. Comput. Phys.* **1997**, *131* (1), 233–240.
- (37) Elber, R. Calculation of the Potential of Mean Force using Molecular Dynamics with Linear Constraints: An Application to a Conformational Transition in a Solvated Dipeptide. *J. Chem. Phys.* **1990**, *93* (6), 4312–4321.
- (38) Andersen, H. C. Rattle: A “Velocity” Version of the SHAKE Algorithm for Molecular Dynamics Calculations. *J. Comput. Phys.* **1983**, *52* (1), 24–34.
- (39) Lennox, C.; Chadwick, M. *Mathematics for Engineers and Applied Scientists*, 2nd ed.; Heinemann Educational Books Ltd: London, 1977.
- (40) Sadhukhan, S.; Noz, D. M.; Adamo, C.; Scuseria, G. E. Predicting Proton Transfer Barriers with Density Functional Methods. *Chem. Phys. Lett.* **1999**, *306* (1–2), 83–87.
- (41) Pohorille, A.; Jarzynski, C.; Chipot, C. Good Practices in Free-Energy Calculations. *J. Phys. Chem. B* **2010**, *114* (32), 10235–10253.