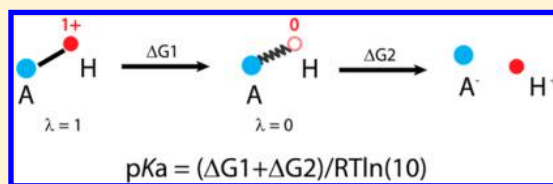


Computing pK_a Values with a Mixing Hamiltonian Quantum Mechanical/Molecular Mechanical Approach

Yang Liu,^{†,||} Xiaoli Fan,^{‡,§,||} Yingdi Jin,[†] Xiangqian Hu,^{*,§} and Hao Hu^{*,†}[†]Department of Chemistry, The University of Hong Kong, Pokfulam Road, Hong Kong[‡]School of Materials Science and Engineering, Northwestern Polytechnical University, Xi'an, Shannxi, 710072, China[§]Department of Chemistry, Duke University, Durham, North Carolina 27708, United States

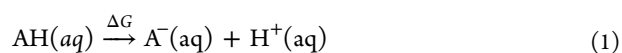
S Supporting Information

ABSTRACT: Accurate computation of the pK_a value of a compound in solution is important but challenging. Here, a new mixing quantum mechanical/molecular mechanical (QM/MM) Hamiltonian method is developed to simulate the free-energy change associated with the protonation/deprotonation processes in solution. The mixing Hamiltonian method is designed for efficient quantum mechanical free-energy simulations by alchemically varying the nuclear potential, i.e., the nuclear charge of the transforming nucleus. In pK_a calculation, the charge on the proton is varied in fraction between 0 and 1, corresponding to the fully deprotonated and protonated states, respectively. Inspired by the mixing potential QM/MM free energy simulation method developed previously [H. Hu and W. T. Yang, *J. Chem. Phys.* **2005**, 123, 041102], this method succeeds many advantages of a large class of λ -coupled free-energy simulation methods and the linear combination of atomic potential approach. Theory and technique details of this method, along with the calculation results of the pK_a of methanol and methanethiol molecules in aqueous solution, are reported. The results show satisfactory agreement with the experimental data.



INTRODUCTION

pK_a is a fundamental property that plays key roles in many chemical and biochemical reactions. Accurate calculation of pK_a value with ab initio quantum mechanics (QM) is very much desired as a means to avoid the technical complications in experiments. The pK_a value of an acid molecule AH is related to the deprotonation free energy (ΔG) of the reaction



through $pK_a = \Delta G / (RT) \ln(10)$, where R is the gas constant and T is the temperature. Therefore, the calculation of pK_a essentially becomes a free-energy simulation problem.

Unlike many other reactions with well-characterized reaction coordinates, the deprotonation process encounters several technical challenges in free-energy simulations. Of many issues to be concerned, one of the biggest difficulties is that the chemical/molecular form of the solvated proton is not well-resolved in both calculations and experiments.¹ With the number of proton-bound water molecules unknown, excessively long simulations are required to provide sufficient sampling of different bonding states. The length of the simulation would be computationally prohibitive if QM methods are applied to the entire simulation system. Even if combined quantum mechanical/molecular mechanical (QM/MM) methods are used to reduce the computational cost, the size of the QM subsystem would have to vary during the simulation to reflect states with different number of water molecules bonding to the proton. Therefore, instead of direct simulation of the proton solvation process, a more feasible approach is to use a reference

value for the contribution of the hydration of proton. This reference value can be taken from other studies or derived from a series of simulations of different compounds with comparison to known experimental pK_a data. Note the former approach is a common practice employed in numerous QM calculations using continuum solvation models.²

Several groups have made important contributions to the QM/MM simulations of pK_a in solution and in enzymes.^{3–6} Most studies avoided the direct simulation of the protonation/deprotonation process. Instead, pK_a was obtained by computing the difference between the solvation free energy of the acid and base molecules. Cui and co-workers made significant progress in direct free energy simulation of the protonation process by using λ -coupled energy function of the acid and base states.⁷ Applications of the method to several enzyme and protein systems have provided useful insights to the protein function and thermodynamics.^{7–10}

Recently, Sprik and co-workers successfully applied the ab initio molecular dynamics (MD) method, i.e., Car–Parrinello molecular dynamics method,¹¹ to simulate the proton and electron transfer process in solutions.^{12–21} Specifically, the pK_a values were computed by density functional theory (DFT)-based MD simulations. The free-energy simulations were also based on thermodynamics integration (TI)²² with the mixing of the energy function of two states, which is a well-developed method in the free-energy simulation of chemical processes.²³

Received: May 16, 2013

Published: July 25, 2013

Here, we propose a new scheme based on the mixing of QM Hamiltonians in QM/MM simulations. We allow the electrostatic potential arisen from the proton to present in partial proportion in the QM or QM/MM Hamiltonian. As a result, there appears to be a fractional proton experienced by other parts of the simulation system. The fraction of the proton can then serve as an order parameter for alchemical free-energy simulations. We applied this method to QM/MM simulations of the pK_a values of methanol and methanethiol in aqueous solution. The simulation method and results are reported.

THEORY

Mixing Hamiltonian for Alchemical Transformation of a Proton. Conventional free-energy simulations often employ a λ -coupled energy function to construct an artificial transformation path between two states.^{23–25} For two thermodynamic states A and B with energy functions E_A and E_B , one can carry out free-energy simulation along a virtual path characterized by a general order parameter λ , as long as the energy function $E(\lambda)$ meets the requirement

$$E(\lambda) = \begin{cases} E_A & \lambda = 0 \\ E_B & \lambda = 1 \end{cases} \quad (2)$$

Note that, depending on the specific methods used to compute the free energies, the energy $E(\lambda)$ does not have to be continuous with respect to the transformation coordinate λ , nor does it have to be an explicit function of λ . There is no restriction on the form of the energy function in this framework; therefore, the QM/MM energy function can be employed very naturally.^{26,27} This is also the foundation for many recent redox and pK_a simulations.^{6,12–20}

When a QM/MM energy function is used, the energy of the system is usually obtained by self-consistently solving the eigen problem for a QM/MM Hamiltonian in the context of Born–Oppenheimer approximation. If the number of electrons is fixed, the wave function of the two end-states and the intermediate states of the transformation process can be expanded with the same local Gaussian basis. This provides the feasibility for transforming the Hamiltonian, instead of the energy of the system.

With the electrostatic embedding scheme, the QM/MM electrostatic Hamiltonians for the two states of the deprotonation process, namely, the acid and the base, are

$$\begin{aligned} H_{\text{ele}}(\text{AH}) = & \frac{-\hbar^2}{2m_e} \sum_{\text{ele}} \nabla^2 - \sum_{\text{ele}} \sum_{i \in A^-} \frac{Z_i}{|\mathbf{r} - \mathbf{R}_i|} \\ & - \sum_{\text{ele}} \frac{Z_H}{|\mathbf{r} - \mathbf{R}_H|} + \sum_{\text{ele}} \sum_{\text{ele}} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \\ & + \sum_{\substack{i,j \in A^- \\ i \neq j}} \frac{Z_i Z_j}{|\mathbf{R}_j - \mathbf{R}_i|} + \sum_{i \in A^-} \frac{Z_i Z_H}{|\mathbf{R}_H - \mathbf{R}_i|} \\ & + \sum_{\substack{i \in A^- \\ j \in \text{MM}}} \frac{Z_i q_j}{|\mathbf{R}_j - \mathbf{R}_i|} + \sum_{j \in \text{MM}} \frac{Z_H q_j}{|\mathbf{R}_j - \mathbf{R}_H|} \\ & + \sum_{\substack{i,j \in \text{MM} \\ i \neq j}} \frac{q_i q_j}{|\mathbf{R}_j - \mathbf{R}_i|} \end{aligned} \quad (3)$$

and

$$\begin{aligned} H_{\text{ele}}(A^-) = & \frac{-\hbar^2}{2m_e} \sum_{\text{ele}} \nabla^2 - \sum_{\text{ele}} \sum_{i \in A^-} \frac{Z_i}{|\mathbf{r} - \mathbf{R}_i|} \\ & + \sum_{\text{ele}} \sum_{\text{ele}} \frac{1}{|\mathbf{r} - \mathbf{r}'|} + \sum_{\substack{i,j \in A^- \\ i \neq j}} \frac{Z_i Z_j}{|\mathbf{R}_j - \mathbf{R}_i|} \\ & + \sum_{\substack{i \in A^- \\ j \in \text{MM}}} \frac{Z_i q_j}{|\mathbf{R}_j - \mathbf{R}_i|} + \sum_{\substack{i,j \in \text{MM} \\ i \neq j}} \frac{q_i q_j}{|\mathbf{R}_j - \mathbf{R}_i|} \end{aligned} \quad (4)$$

respectively. Here, Z_H is the nuclear charge of the disappearing proton, whose value should be one atomic unit (a.u.). Z_i is the nuclear charge of any other QM atom i , and q_j is the point charge of an MM atom j . \mathbf{r} is the position of electrons, and \mathbf{R} is the position of QM or MM atoms. Now, one may construct a λ -mixed transformation Hamiltonian as

$$\begin{aligned} H_{\text{ele}}(\lambda) = & \frac{-\hbar^2}{2m_e} \sum_{\text{ele}} \nabla^2 - \sum_{\text{ele}} \sum_{i \in A^-} \frac{Z_i}{|\mathbf{r} - \mathbf{R}_i|} \\ & - \sum_{\text{ele}} \frac{\lambda Z_H}{|\mathbf{r} - \mathbf{R}_H|} + \sum_{\text{ele}} \sum_{\text{ele}} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \\ & + \sum_{\substack{i,j \in A^- \\ i \neq j}} \frac{Z_i Z_j}{|\mathbf{R}_j - \mathbf{R}_i|} + \sum_{i \in A^-} \frac{\lambda Z_i Z_H}{|\mathbf{R}_H - \mathbf{R}_i|} \\ & + \sum_{\substack{i \in A^- \\ j \in \text{MM}}} \frac{Z_i q_j}{|\mathbf{R}_j - \mathbf{R}_i|} + \sum_{j \in \text{MM}} \frac{\lambda Z_H q_j}{|\mathbf{R}_j - \mathbf{R}_H|} \\ & + \sum_{\substack{i,j \in \text{MM} \\ i \neq j}} \frac{q_i q_j}{|\mathbf{R}_j - \mathbf{R}_i|} \end{aligned} \quad (5)$$

It immediately shows that

$$H_{\text{ele}}(\lambda) = \begin{cases} H_{\text{ele}}(\text{AH}) & \lambda = 1 \\ H_{\text{ele}}(A^-) & \lambda = 0 \end{cases} \quad (6)$$

and subsequently

$$E_{\text{ele}}(\lambda) = \begin{cases} E_{\text{ele}}(\text{AH}) & \lambda = 1 \\ E_{\text{ele}}(A^-) & \lambda = 0 \end{cases} \quad (7)$$

Since the λ -coupled QM/MM electrostatic Hamiltonian meets the requirement specified in eq 2, it can be used to simulate the free-energy change along the path defined by λ . Similar to free-energy simulation with MM force field, nonlinear coupling of λ is also possible as long as eq 6 is satisfied.

When the Hamiltonian is changing, the energy derivative with respect to λ may be computed with the Hellmann–Feynman theorem as

$$\begin{aligned}
\frac{\partial E_{\text{ele}}(\lambda)}{\partial \lambda} &= \frac{\partial \langle \Psi | H_{\text{ele}}(\lambda) | \Psi \rangle}{\partial \lambda} \\
&= \left\langle \Psi \left| \frac{-Z_{\text{H}}}{|\mathbf{r} - \mathbf{R}_{\text{H}}|} + \sum_{i \in \text{A}^-} \frac{Z_i Z_{\text{H}}}{|\mathbf{R}_{\text{H}} - \mathbf{R}_i|} \right. \right. \\
&\quad \left. \left. + \sum_{j \in \text{MM}} \frac{Z_{\text{H}} q_j}{|\mathbf{R}_j - \mathbf{R}_{\text{H}}|} \right| \Psi \right\rangle \\
&= \left\langle \Psi \left| \frac{-Z_{\text{H}}}{|\mathbf{r} - \mathbf{R}_{\text{H}}|} \right| \Psi \right\rangle + \sum_{i \in \text{A}^-} \frac{Z_i Z_{\text{H}}}{|\mathbf{R}_{\text{H}} - \mathbf{R}_i|} \\
&\quad + \sum_{j \in \text{MM}} \frac{Z_{\text{H}} q_j}{|\mathbf{R}_j - \mathbf{R}_{\text{H}}|} \\
&= Z_{\text{H}} \left(\left\langle \Psi \left| \frac{-1}{|\mathbf{r} - \mathbf{R}_{\text{H}}|} \right| \Psi \right\rangle + \sum_{i \in \text{A}^-} \frac{Z_i}{|\mathbf{R}_{\text{H}} - \mathbf{R}_i|} \right. \\
&\quad \left. + \sum_{j \in \text{MM}} \frac{q_j}{|\mathbf{R}_j - \mathbf{R}_{\text{H}}|} \right) \quad (8)
\end{aligned}$$

In the conventional Hartree–Fock or DFT methods, the three terms in the parentheses on the right-hand side are the electrostatic potential (ESP) at the position of the proton from different sources. Specifically, these are the ESP from the QM electrons, from all QM nuclei except the disappearing proton, and from the MM (point) charges, respectively. The sum of the three terms is the total ESP on the proton. In other words, the energy gradient for the process of linear transformation of the QM proton, between fully present in the acid and fully disappeared in the base, is the total electrostatic potential experienced by the proton multiplied by a proton nuclear charge Z_{H} . This property has been used to predict the $\text{p}K_{\text{a}}$ of small molecules.²⁸ The availability of the energy gradient in such a simple form allows one to compute the electrostatic contribution of the free energy with the TI method²² as

$$\Delta G_{\text{ele}} = \int_1^0 \left\langle \frac{\partial E_{\text{ele}}(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (9)$$

Computing the Deprotonation Free Energy. The total energy of an acid molecule in solution described by QM/MM methods generally also contains van der Waals (vdW) interactions between QM and MM subsystems, such as

$$\begin{aligned}
E_{\text{AH}} &= \langle \Psi | H_{\text{ele}}(\text{HA}) | \Psi \rangle + \sum_{\substack{i \in \text{A}^- \\ j \in \text{MM}}} E_{\text{vdW}}(R_{ij}) \\
&\quad + \sum_{j \in \text{MM}} E_{\text{vdW}}(R_{\text{H}j}) + E_{\text{MM}} \quad (10)
\end{aligned}$$

where $E_{\text{vdW}}(R_{ij})$ and $E_{\text{vdW}}(R_{\text{H}j})$ are vdW energy between QM atom i and MM atom j , and between the disappearing QM proton and MM atom j , respectively. Clearly, the vdW interactions between the transforming proton and the rest of the MM molecular system contribute to the total free-energy change. In the common QM/MM methods, the pair of conjugate acid and base molecules often has different vdW interactions with the solvents. This issue must also be taken into account in the current work. This simple QM/MM scheme with different fixed vdW parameters for two molecular forms does not consider the fact that, in principle, QM atoms should have polarized, different vdW interactions when the QM

molecule changes from acid to base.²⁹ The part of the vdW contribution can be computed either through a separated process, or being coupled to the electrostatic interactions.

With the approach of λ -mixed Hamiltonian, one can effectively remove the proton from the acid in solution and transfer it to vacuum, while the energy of a free proton in vacuum is zero in the Born–Oppenheimer framework. The total free energy for $\text{p}K_{\text{a}}$ calculation then requires the free energy of a second process, which solvates the proton back into solution (Figure 1). The total free energy is then

$$\Delta G = \Delta G_{\text{dp}} + \Delta G_{\text{solv}}^{\text{H}^+} \quad (11)$$

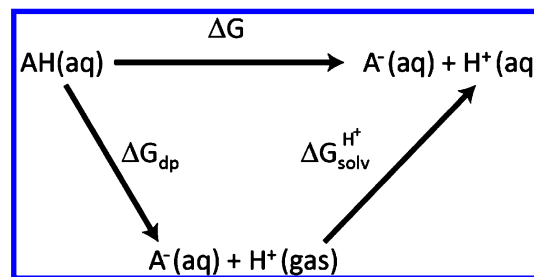


Figure 1. Thermodynamic cycle for $\text{p}K_{\text{a}}$ simulation.

Direct computation of the proton solvation free energy suffers two notorious technical difficulties. The first problem is the aforementioned unknown number of water molecules bound to the proton in solution. The second problem is the free-energy cost for the proton to cross the vacuum/liquid interface. This is a term that often causes confusion in simulations, because it is difficult to model accurately.^{30–32} In the current work, we chose not to directly tackle these issues; instead, we use existing experimental/theoretical data of proton solvation to complete the thermodynamic cycle in Figure 1.

Strategy for Treating the Disappearing Proton. One practical issue needs to be carefully addressed in the current QM/MM or QM free-energy simulation with the approach of scaling proton charge. That is, the bonding interaction between the changing proton and the cognate heavy atom becomes weaker and weaker when the proton charge is scaled down. The weakening of the interaction will reach to a point that the thermal energy is sufficiently large to break the covalent bond. Once this happens, the reforming of the bond is very unlikely to be observed within the limited time scale of the QM/MM or QM simulations. The consequence would be very large fluctuations in the calculation of energy and energy gradient in free-energy simulations.

Similar to previous work,²⁰ we employed a harmonic restraint on the distance between the disappearing proton and the cognate heavy atom. We observed in simulations that the bonding interactions are usually strong enough to hold the two atoms together when $\lambda \geq 0.3$. From the simulation data for $\lambda \geq 0.3$, a harmonic potential was designed and applied to the system when $\lambda < 0.3$. The formula of the restraining potential for methanol and methanethiol are, respectively,

$$\begin{aligned}
V_{\text{rstr}}(r, \lambda) &= 0.5 \times 128(\lambda - 0.3)^2 \\
&\quad \times \left[r - (0.94852 + 0.24869 \exp\left(-\frac{\lambda}{0.55083}\right)) \right]^2 \quad (12)
\end{aligned}$$

and

$$V_{\text{rstr}}(r, \lambda) = 0.5 \times 128(\lambda - 0.3)^2 \times \left[r - (1.28807 + 0.53038 \exp\left(-\frac{\lambda}{0.47855}\right)) \right]^2 \quad (13)$$

Here, r is the distance between the proton and its cognate heavy atom. Because the restraining potential is λ -dependent, it contributes to the computed free energy.

Complete Transformation Path. With the vdW contributions considered separately from the electrostatic interaction, the completed transformation path now constitutes the following steps:

(i) *Scaling down linearly the nuclear charge of the proton from 1.0 a.u. to 0.3 a.u. with a step size of 0.1.* The vdW parameters are kept unchanged. The free-energy change is computed by TI. That is, the ESP at the position of the proton is computed as the derivative of energy with respect to λ (eq 9). At the end of this step, the proton still possesses a fractional charge of 0.3 a.u. and full vdW interactions with the rest of the simulation system.

(ii) *Scaling down the nuclear charge of the proton from 0.3 a.u. to 0.0 a.u.* For each given value of $\lambda = 0.2, 0.1, 0.05, 0.01$ and 0.0, a harmonic restraining potential (eq 12 or 13) is applied to maintain the O–H or S–H bond. The equilibrium distance is derived from a curve fitted from simulations in step (i) for λ between 1.0 and 0.3. The energy derivative of λ now contains two components (see eq 15 below): one from the ESP at the proton position (eq 8) and another from the restraining potential (eq 12). At the end, the proton has no charge, but still carries its bases for QM calculations and possesses full vdW interactions. It is linked to the cognate heavy atom with a classical restraining potential. This is in fact a base molecule connected to a dummy proton, denoted as AD[−].

(iii) *Computing the free-energy contribution from the vdW interactions between the QM and MM subsystems.* At the end point, the proton has no charge and vdW interactions to the rest of the system, but it is still attached to the heavy atom and carries its bases. The free-energy component is computed by the free-energy perturbation (FEP) method:³³

$$\Delta G_{\text{vdW}} = -k_{\text{B}}T \times \ln \left\langle \exp \left[-\frac{1}{k_{\text{B}}T} (E_{\text{vdW,A}^-}(\mathbf{R}_{\text{AD}^-}) - E_{\text{vdW,AH}}(\mathbf{R}_{\text{AD}^-})) \right] \right\rangle_{\text{AD}^-} \quad (14)$$

Here, $E_{\text{vdW,A}^-}(\mathbf{R}_{\text{AD}^-})$ and $E_{\text{vdW,AH}}(\mathbf{R}_{\text{AD}^-})$ are the vdW energies of the dummy-proton bound base (AD[−]) computed using the vdW parameters of the base and acid, respectively.

(iv) *Removing the contribution of the extra bases carried by the dummy proton.* This free-energy term is also computed using the free-energy perturbation technique.

It is certainly possible to combine step (iii) into steps (i) and (ii), but we leave it here as a separate step for the sake of clarification.

Direct Contribution of Restraining Potential to Thermodynamic Integration. When the proton charge is scaled down, a restraining potential V_{rstr} is applied to keep the proton bonded to the conjugate base. The restraining potential will make a direct contribution to the free energy via a component in the total energy gradient as

$$\frac{\partial \Delta G_{\text{ele}}}{\partial \lambda} = \left\langle V_{\text{ESP}}(\lambda) + \frac{\partial V_{\text{rstr}}(r, \lambda)}{\partial \lambda} \right\rangle_{\lambda} \quad (15)$$

This term is included in the calculation of ΔG_{ele} .

Free-Energy Correction to the Restrained State.

According to the thermodynamic cycle (Figure 1), the proton in the end state should be a free particle in the gas phase. This is different from the end state of the simulation, where the proton is still bound to the cognate heavy atom with the restraining potential. Thus, there is a free-energy difference between the two states.

$$\Delta G_{\text{rstr}} = -RT \ln \left(\frac{q_{\text{A}^-} q_{\text{H}_{\text{gas}}}}{q_{\text{AH}_{\text{rstr}}}} \right) \quad (16)$$

$q_{\text{AH}_{\text{rstr}}}$, q_{A^-} , and $q_{\text{H}_{\text{gas}}}$ are the partition functions of the base with proton attached by a restraining potential, base molecule, and proton in gas phase, respectively. As the proton in this state has no physical interactions with the rest of the atoms of the base molecule, the restraint applied to the proton only provides three additional modes of motion for the proton: bond vibration, bond angle bending, and rotation. Therefore, we further approximate

$$q_{\text{AH}_{\text{rstr}}} = q_{\text{A}^-} q_{\text{H}_{\text{rstr}}} \quad (17)$$

The free-energy correction becomes

$$\Delta G_{\text{rstr}} = -RT \ln \left(\frac{q_{\text{H}_{\text{gas}}}}{q_{\text{H}_{\text{rstr}}}} \right) = G_{\text{H}_{\text{gas}}} - G_{\text{H}_{\text{rstr}}} \quad (18)$$

$q_{\text{H}_{\text{rstr}}}$ will be computed as a configurational integral in Cartesian space with the distribution of three components, i.e., the O–H bond, the C–O–H bond angle, and the H–C–O–H dihedral angle, determined from QM/MM MD simulations in the restrained state. $q_{\text{H}_{\text{gas}}}$ is the partition function of an ideal gas particle, which can be easily evaluated using a QM particle-in-a-box model.

One issue for the calculation of the free-energy contribution of the restraint requires special attention here. That is, the standard state for eq 16 is 1 mol/L instead of 1 bar. As clarified in previous theoretical analysis,³⁴ this can be corrected by using −264.0 kcal/mol for the proton solvation free energy, instead of −265.9 kcal/mol.³⁵

Free-Energy Correction to Ghost Basis. When the QM methods use localized atom-centered basis sets such as Gaussian or Slater functions, the proton always carries a finite basis, even if it has no physical interactions with the rest. This effectively leads to a situation where the base molecule is simulated with additional bases from the disappeared proton. It is certainly debatable whether the contribution of the extra bases should be removed, since, in principle, larger bases always provide better QM energies. This is a unique issue for QM methods using local basis sets. There should not be such a problem if nonlocal plane waves are used. To make the results consistent with customary gas-phase calculations, we remove the contribution of the additional bases from the disappearing ghost proton using the FEP method.³³

$$\Delta G_{\text{ghost basis}} = -k_{\text{B}}T \times \ln \left\langle \exp \left[-\frac{1}{k_{\text{B}}T} (E_{\text{QM}}(\text{AH}, \mathbf{R}_{\text{AD}^-}) - E_{\text{QM}}(\text{A}^-, \mathbf{R}_{\text{AD}^-})) \right] \right\rangle_{\text{AD}^-} \quad (19)$$

Here, $E_{\text{QM}}(\text{AH}, \mathbf{R}_{\text{AD}}^-)$ and $E_{\text{QM}}(\text{A}^-, \mathbf{R}_{\text{AD}}^-)$ are the QM energy of the base molecule computed with and without the bases of the diminishing proton, respectively.

Quantum Correction to Nuclear Motions. Up to this point, both the acid and base are described in the context of Newtonian dynamics of nuclei. It is known that the quantum corrections from the nuclear motions make important contributions to the free-energy difference.²⁰ The quantum corrections include two components. One is the difference between the partition function described by the quantum mechanical nuclear motions and by the classical Newtonian mechanics. The other is the zero point energy that arises from the quantum mechanical nuclear motions. The latter contributions will be included in the next section where the systematic error of specific QM methods is evaluated. The former term will be computed using the same technique proposed previously.²⁰ That is, the difference can be evaluated as

$$\Delta G_{\text{qc}} = G_{\text{QM}}^{\text{vib}}(\text{AH}) - G_{\text{classical}}^{\text{vib}}(\text{AH}) \quad (20)$$

where $G_{\text{QM}}^{\text{vib}}(\text{AH})$ and $G_{\text{classical}}^{\text{vib}}(\text{AH})$ are the free energies originated from the nuclear motions described by quantum mechanics and by classical Newtonian mechanics, respectively. Using a simplified vibrational model for the three degrees of freedom involving the proton, we have²⁰

$$G_{\text{QM}}^{\text{vib}}(\text{AH}) = -k_{\text{B}}T \ln \prod_{i=1}^3 \left[\frac{\exp\left(-\frac{h\nu_{\text{AH},i}}{2k_{\text{B}}T}\right)}{1 - \exp\left(-\frac{h\nu_{\text{AH},i}}{k_{\text{B}}T}\right)} \right] \quad (21)$$

$$G_{\text{classical}}^{\text{vib}}(\text{AH}) = -k_{\text{B}}T \ln \prod_{i=1}^3 \left(\frac{k_{\text{B}}T}{h\nu_{\text{AH},i}} \right) \quad (22)$$

where $\nu_{\text{AH},i}$ denotes the frequency of the vibrational mode i of the proton. Note that the rotation has a degeneracy number of 3 for the molecules studied here, but the corresponding contribution will cancel out in the free-energy difference as long as the acid molecule possesses the same structural degeneracy in both the QM and MM energy functions.

Correction of Systematic Errors of QM Methods. The QM method used in the current work is B3LYP/6-31+G*. The use of this specific level of QM theory certainly brings systematic errors into the energy calculation of each molecular state. The correction to the systematic error of the QM method can be made by comparison to experimental data or high accuracy QM calculations in the gas phase. The gas-phase proton affinity (PA) can be computed theoretically as

$$\text{PA}_{\text{calc}} = -\Delta E_{\text{ele}} - \Delta \text{ZPE} - \Delta E_{\text{vib}'} + \frac{5}{2}RT \quad (23)$$

The difference between the calculated and experimental values will be used as a systematic correction to the specific QM methods used in our calculations. That is,

$$\Delta G_{\text{QM}} = \text{PA}_{\text{expr}} - \text{PA}_{\text{calc}} \quad (24)$$

Total Deprotonation Free Energy. The total deprotonation free energy thus contains contributions from individual components as

$$\begin{aligned} \Delta G_{\text{dp}} = & \Delta G_{\text{ele}} + \Delta G_{\text{vdW}} + \Delta G_{\text{rstr}} + \Delta G_{\text{ghost basis}} - \Delta G_{\text{qc}} \\ & + \Delta G_{\text{QM}} \end{aligned} \quad (25)$$

They correspond to the various contributions discussed previously. The first two terms on the right-hand side correspond to the electrostatic (including the restraining potential) and vdW contributions computed in steps (i), (ii), and (iii). The third term is the correction to the restrained state of the proton, as compared to an ideal gas particle in standard state. The fourth term is the correction of the bases carried by the disappeared dummy proton. The last term is the correction of systematic errors of specific QM methods.

SIMULATION DETAILS

All simulations were carried out with an in-house QM4D³⁶ simulation package in which the current method was implemented. The solute molecules were described by B3LYP/6-31+G*,^{37,38} while the solvent water molecules were described by the TIP3P model.³⁹ A cubic solvation box of 48 Å × 48 Å × 48 Å was used for both systems. The CHARMM force field was used to provide the vdW interactions between the QM and MM subsystems.⁴⁰ The MD integration time step was 1 fs. A cutoff of 10 Å was used to evaluate nonbonding interactions in MD simulations. The nonbonding pair list was updated every 8 fs. The mass of the transforming proton was set to 10 times of the normal mass. The increase of the mass of proton is important to ensure the proper thermalization of all solute degrees of freedom when a rigid water model is used in MD simulations.⁴¹ The temperature of the system was maintained at 298 K using a Berendsen thermostat.⁴²

As an illustration of the method, the long-range electrostatic interactions were not considered in MD sampling. For the small size of the molecules simulated here, use of a cutoff of 10 Å seems sufficient to provide correct structure of liquid states. After the MD simulations, the long-range electrostatic contributions were partially recovered through replaying MD trajectories using a nonbonded cutoff of 20 Å to compute all terms used in free-energy calculations. This posterior approach has been employed before in the calculation of redox potential in solution.^{43,44}

For each intermediate transformation state in the free-energy simulations, 40-ps MD samplings were carried out. Four independent samplings were performed to estimate the statistical errors. The conformations were saved every 8 fs and used lately for FEP or TI calculations.

RESULTS AND DISCUSSION

Gas-Phase Properties. To examine whether the molecules with a fractional proton are stable, we ran energy minimizations for the methanol molecule with varied fractions of the proton potential. The bond distance between the disappearing proton and its cognate atom is shown in Figure 2. It appears that the disappearing proton can maintain a stable position without collision to other nuclei or dissociating even when it has a zero nuclear charge. This phenomenon may look surprising at the first sight, but one must understand that it is the result of the Gaussian bases carried by the proton. This conclusion is supported by the observation that calculations with different basis sets show different correlations between the bond length and the proton charge. The proton can adopt any position in space if the bases of the ghost proton were removed or nonlocal basis sets such as plane-waves were used. This bonding

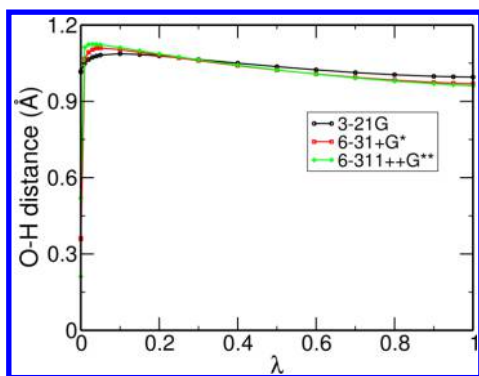


Figure 2. Correlation between minimized O–H distance and the fraction of proton charge in methanol.

behavior between the ghost proton and the basic molecule could be advantageous because it allows MD sampling of this end state. The contribution from the additional bases of the ghost proton could be easily removed by free-energy perturbations, as discussed previously.

Free-Energy Simulation Results. The O–H bond distance and the C–O–H angle in QM/MM MD simulations of methanol are respectively plotted as a function of the fraction of proton (λ) in Figures 3 and 4) When the fraction of the

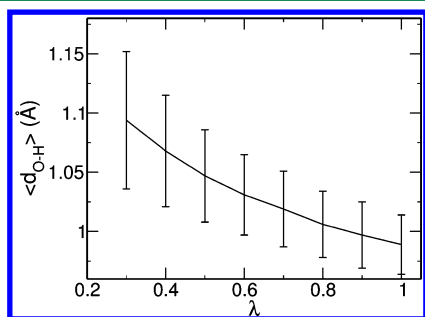


Figure 3. Mean and root-mean-square (rms) fluctuation of O–H bond distance in QM/MM MD simulations of methanol with different fractions of proton.

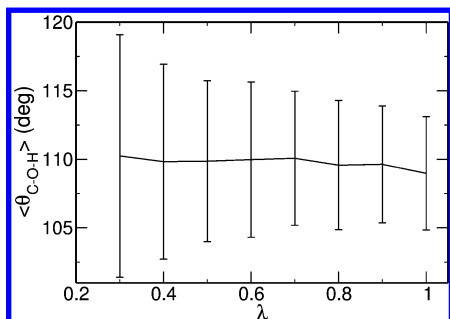


Figure 4. Mean and rms fluctuation of the C–O–H bond angle in QM/MM MD simulations of methanol with different fractions of proton.

proton decreases from 1 to 0.3, the average O–H bond length increases. The magnitude of the fluctuation changes slightly, suggesting that the bonding strength is still significant even for the state with only 30% of the proton potential.

The gradient (eq 8) computed for individual simulation window and the total contribution of the electrostatic and restraint component of the deprotonation process are reported

in Tables S1 and S2 in the Supporting Information. Quantum corrections to the individual mode (eq 20) are shown in Table 1. Because of the weaker bonding interactions and heavier mass

Table 1. Free-Energy Corrections to the Classical Nuclear Motions in Methanol and Methanethiol^a

methanol		methanethiol	
$\nu_{\text{AH},i}$ (cm^{-1})	$\Delta G_{\text{qc},i}$ (kcal/mol)	$\nu_{\text{AH},i}$ (cm^{-1})	$\Delta G_{\text{qc},i}$ (kcal/mol)
325	0.06	254	0.04
1370	0.84	805	0.34
3764	3.66	2681	2.32

$$\Delta G_{\text{qc}} = 4.56$$

$$\Delta G_{\text{qc}} = 2.70$$

^aThe free-energy difference for each vibration mode was computed using eqs 20–22.

of the sulfur atom, the magnitude of the corrections of methanethiol is smaller than that of methanol. The systematic correction to the QM methods used in this study (eq 24) is reported in Table 2. The magnitude of the correction is larger

Table 2. Free-Energy Corrections to the QM Method

	methanol	methanethiol
$\Delta E_{\text{ele}} + \Delta \text{ZPE}$ (kcal/mol)	−374.9	−352.2
ΔE_{vib} (kcal/mol)	0.3	0.3
PA_{expr} ^a (kcal/mol)	379.302	358.987
PA_{calc} (kcal/mol)	376.2	353.4
ΔG_{QM} (kcal/mol)	−3.1	−5.6

^aData taken from ref 58.

for methanethiol than for methanol, presumably due to the 6-31+G* basis set used here. The contribution to the restrained phase space sampling is shown in Table 3.

Table 3. Free-Energy Corrections to the Phase Space Sampling of the Restraining Potential

	methanol	methanethiol
$\Delta G_{\text{H}_{\text{gs}}}$ (kcal/mol)	−6.3	−6.3
$\Delta G_{\text{H}_{\text{str}}}$ (kcal/mol)	0.9	1.3
ΔG_{rstr} (kcal/mol)	−5.4	−5.0

The total and individual components of the deprotonation free energy of methanol and methanethiol are shown in Table 4. Using a proton solvation free energy of −264.0 kcal/mol,^{34,35} the computed pK_{a} value of methanol and methanethiol are in good agreement with the experimental data.

Analysis of Sources of Error. We group the systematic errors into two categories: those associated with the QM/MM energy function, and those associated with classical Newtonian description of nuclear motions in phase space sampling.

The first and foremost factor to affect the simulation accuracy is the intrinsic accuracy of the specific quantum chemistry methods. Among many QM methods, density functional theory is popular in chemistry because it strikes a balance between calculation accuracy and computation speed. Nevertheless, DFT suffers from well-known deficiencies, such as lacking of native dispersions and prone of delocalization error,^{46,47} etc. Furthermore, like others, the error associated with the use of a specific, limited-size basis set is also intrinsic to all DFT-based methods. The large part of the error from the

Table 4. Total Deprotonation Free Energy and pK_a Values^a

	methanol	methanethiol
ΔG_{ele} (kcal/mol)	303.6 ± 0.8	286.6 ± 0.6
ΔG_{vdW} (kcal/mol)	-3.4 ± 0.1	4.9 ± 0.1
ΔG_{rstr} (kcal/mol)	-5.4	-5
$\Delta G_{\text{ghost basis}}$ (kcal/mol)	-1.10 ± 0.03	-1.43 ± 0.04
ΔG_{qc} (kcal/mol)	4.6	2.7
ΔG_{QM} (kcal/mol)	-3.1	-5.6
$\Delta G_{\text{solv}}^{\text{H}^+}$ (kcal/mol) ^b	-264.0	-264.0
ΔG (kcal/mol)	22.0 ± 0.9	12.8 ± 0.7
pK_a (calc)	16.1 ± 0.7	9.4 ± 0.5
pK_a (expr) ^c	15.54 ± 0.7	10.4 ± 0.6
ΔpK_a	0.6	-1.0

^aThe root-mean-square deviation (RMSD) was separately estimated for each term from four parallel simulations. ^bData taken from refs 34 and 35. ^cData taken from ref 59.

employment of specific DFT methods and basis sets can be corrected by analyzing calculation results in the gas phase. However, currently, the effects from multibody effects in condensed phases are still lacking.

Compared to the full QM description of the entire molecular system, partition and implementation of QM/MM schemes further lead to several known deficiencies in accuracy. Not surprisingly, most problems can be blamed to the unbalanced treatment of electronic degrees of freedom in QM and MM regions. Unlike QM, the MM subsystem usually uses simple fixed point charges to represent electrostatic interactions. The charges are often prepolarized to compensate for the effect of polarization. The overpolarized point charges generally will overstabilize polar molecules in solution. The degree of overstabilization will be different between the acid and base molecules, thus leading to a systematic error for ΔG_{ele} . One possible solution to this problem might be fitting different sets of QM/MM vdW parameters for the acid and base molecules, respectively, but this is currently unavailable.

The effects of QM/MM vdW interactions on QM/MM pK_a calculations are important but complicated. First, the difference of QM/MM vdW interactions in different protonation states makes a direct contribution to the total free-energy difference ΔG_{vdW} . Second, the QM/MM vdW interactions will affect the spatial distribution of the solvent molecules around the QM solute, thus contribute to the electrostatic component of free energy ΔG_{ele} . Both contributions highlight the importance of accurate QM/MM vdW interactions, as discovered previously.⁴⁸ We also like to point out that the combination of MM point charge and classical vdW interactions leads to intrinsic deficiencies in the computed electron density, dipole moment, and other properties of the QM subsystem,^{49,50} thus limiting the accuracy of QM/MM free-energy simulations.

Long-range QM/MM electrostatic interactions make important contributions to the deprotonation free energy, but were not rigorously considered in the current work. Since a large cutoff was used in the evaluation of nonbonding interactions, the error should not be very significant from our previous experience in modeling redox reactions.^{43,44} Nonetheless, we recently developed a QM/MM particle-mesh Ewald scheme to compute the long-range QM/MM electrostatic interactions. Simulations are undergoing to treat this properly.

The error associated with the classical nuclear motions is slightly complicated. Full QM simulation of nuclear motions of

molecules of more than 10 atoms is currently prohibitive. The only viable approach at this time is to apply simplified models to approximate the real motions. When the motional modes associated with the proton are distinct from other motions, and so the number of relevant motional modes is small, simple models such as the quantum harmonic oscillator can be applied to provide approximate description of the proton motions. This is the assumption undertaken in the current and previous research.

There is yet one term whose contribution is unclear at this time. In the current scheme, the contribution of the bases carried by the ghost proton was removed by FEP. This treatment is consistent with the later calculation of the systematic error of the specific quantum mechanical methods. Whether or not a more-sophisticated approach such as basis set superposition error should be conducted is a question remains to be examined.

Further Extension. The mixing Hamiltonian method reported here roots deeply into the classical λ -coupled free-energy simulation approach.^{23–25,51} Instead of mixing the energy function of two states, it mixes the Hamiltonian of two states, thus reducing the cost of QM calculations to half of that of the methods using mixed energy function. The method also shares a similar spirit to the two methods developed previously. The first is the linear combination of atomic potential approach developed for the inverse de novo design of new molecules for special properties.^{52–57} The other is a QM/MM free-energy simulation approach in which the external potential of two different states were mixed to give a λ -coupled QM Hamiltonian.²⁶ The latter method, in combination with the current method, forms a complete QM/MM free-energy simulation scheme using a mixed Hamiltonian.

The approach proposed here can be easily extended to cases where nuclei other than the proton are varied, as long as the number of electrons is conserved. When the target process compresses changes in both nuclear type and electron numbers, the current scheme would require different sets of simulations to separately model the changes of the nuclear type and number of electrons.^{43,44}

CONCLUSION

Here, we have reported a new mixing Hamiltonian approach for the QM/MM free-energy simulation of pK_a in condensed phases. The method was applied to the calculation of pK_a of methanol and methanethiol in aqueous solution and show satisfactory results, compared to experiments. The method can be easily extended to other processes in which the number of electrons is fixed; thus, it has great application potential in the simulation of many important biochemical processes.

ASSOCIATED CONTENT

Supporting Information

Two tables for contribution of electrostatic and restraining potential to the deprotonation free energy of methanol and methanethiol. One figure for the average number of SCF iterations in MD simulation of different λ states. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: haohu@hku.hk.

Author Contributions

[†]The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the Research Grants Council of Hong Kong, the University Development Fund on Fast Algorithms, Strategic Theme on Computational Sciences, and Seed Funding for Basic Research at the University of Hong Kong for providing financial supports, the high-performance computing facility of the computer center at HKU for providing computing resources. We thank one referee for pointing out the issue of standard state in making corrections to the restrained state of the proton.

ABBREVIATIONS

DFT, density functional theory; ESP, electrostatic potential; FEP, free-energy perturbation; MD, molecular dynamics; MM, molecular mechanics; QM, quantum mechanics; QM/MM, quantum mechanical/molecular mechanical; TI, thermodynamic integration; vdW, van der Waals

REFERENCES

- (1) Marx, D. Proton transfer 200 years after von Grotthuss: Insights from ab initio simulations. *ChemPhysChem* **2006**, *7*, 1848.
- (2) Mennucci, B.; Cammi, R., *Continuum Solvation Models in Chemical Physics: From Theory to Applications*; John Wiley & Sons, Ltd.: West Sussex, England, 2007.
- (3) Warshel, A. Calculations of Enzymatic Reactions: Calculations of pK_a, Proton Transfer Reactions, and General Acid Catalysis Reactions in Enzymes. *Biochemistry* **1981**, *20*, 3167.
- (4) Jorgensen, W. L.; Briggs, J. M.; Gao, J. A Priori Calculations of pK_a's for Organic Compounds in Water. The pK_a of Ethane. *J. Am. Chem. Soc.* **1987**, *109*, 6857.
- (5) Gao, J.; Li, N.; Freindorf, M. Hybrid QM/MM Simulations Yield the Ground and Excited State pK_a Difference: Phenol in Aqueous Solution. *J. Am. Chem. Soc.* **1996**, *118*, 4912.
- (6) Li, G. H.; Cui, Q. pK(a) calculations with QM/MM free energy perturbations. *J. Phys. Chem. B* **2003**, *107*, 14521.
- (7) Riccardi, D.; Schaefer, P.; Cui, Q. pK(a) calculations in solution and proteins with QM/MM free energy perturbation simulations: A quantitative test of QM/MM protocols. *J. Phys. Chem. B* **2005**, *109*, 17715.
- (8) Ghosh, N.; Cui, Q. pK(a) of residue 66 in Staphylococcal nuclease. I. Insights from QM/MM simulations with conventional sampling. *J. Phys. Chem. B* **2008**, *112*, 8387.
- (9) Ghosh, N.; Prat-Resina, X.; Gunner, M. R.; Cui, Q. Microscopic pK(a) Analysis of Glu286 in Cytochrome c Oxidase (*Rhodobacter sphaeroides*): Toward a Calibrated Molecular Model. *Biochemistry* **2009**, *48*, 2468.
- (10) Goyal, P.; Ghosh, N.; Phatak, P.; Clemens, M.; Gaus, M.; Elstner, M.; Cui, Q. Proton Storage Site in Bacteriorhodopsin: New Insights from Quantum Mechanics/Molecular Mechanics Simulations of Microscopic pK(a) and Infrared Spectra. *J. Am. Chem. Soc.* **2011**, *133*, 14981.
- (11) Car, R.; Parrinello, M. Unified approach for molecular-dynamics and density-functional theory. *Phys. Rev. Lett.* **1985**, *55*, 2471.
- (12) Blumberger, J.; Bernasconi, L.; Tavernelli, I.; Vuilleumier, R.; Sprik, M. Electronic structure and solvation of copper and silver ions: A theoretical picture of a model aqueous redox reaction. *J. Am. Chem. Soc.* **2004**, *126*, 3928.
- (13) Blumberger, J.; Sprik, M. Free energy of oxidation of metal aqua ions by an enforced change of coordination. *J. Phys. Chem. B* **2004**, *108*, 6529.
- (14) Blumberger, J.; Sprik, M. Ab initio molecular dynamics simulation of the aqueous Ru²⁺/Ru³⁺ redox reaction: The Marcus perspective. *J. Phys. Chem. B* **2005**, *109*, 6793.
- (15) Blumberger, J.; Tavernelli, I.; Klein, M. L.; Sprik, M. Diabatic free energy curves and coordination fluctuations for the aqueous Ag⁺/Ag²⁺ redox couple: A biased Born–Oppenheimer molecular dynamics investigation. *J. Chem. Phys.* **2006**, *124*, 064507.
- (16) VandeVondele, J.; Lynden-Bell, R.; Meijer, E. J.; Sprik, M. Density functional theory study of tetrathiafulvalene and thianthrene in acetonitrile: Structure, dynamics, and redox properties. *J. Phys. Chem. B* **2006**, *110*, 3614.
- (17) Sulpizi, M.; Sprik, M. Acidity constants from vertical energy gaps: density functional theory based molecular dynamics implementation. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5238.
- (18) Cheng, J.; Sulpizi, M.; Sprik, M. Redox potentials and pK(a) for benzoquinone from density functional theory based molecular dynamics. *J. Chem. Phys.* **2009**, *131*, 154504.
- (19) Sulpizi, M.; Sprik, M. Acidity constants from DFT-based molecular dynamics simulations. *J. Phys. Condens. Matter* **2010**, *22*, 284116.
- (20) Mangold, M.; Rolland, L.; Costanzo, F.; Sprik, M.; Sulpizi, M.; Blumberger, J. Absolute pK(a) Values and Solvation Structure of Amino Acids from Density Functional Based Molecular Dynamics Simulation. *J. Chem. Theory Comput.* **2011**, *7*, 1951.
- (21) Liu, X. D.; Sprik, M.; Cheng, J. Hydration, acidity and metal complexing of polysulfide species: A first principles molecular dynamics study. *Chem. Phys. Lett.* **2013**, *563*, 9.
- (22) Kirkwood, J. G. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **1935**, *3*, 300.
- (23) Beveridge, D. L.; DiCapua, F. M., Free energy via molecular simulation: A primer. In *Computer Simulation of Biomolecular Systems*, van Gunsteren, W. F.; Weiner, P. K., Eds.; ESCOM: Leiden, The Netherlands, 1989; Vol. 1, pp 1.
- (24) Jorgensen, W. L. Free Energy Calculations: A Breakthrough for Modeling Organic Chemistry in Solution. *Acc. Chem. Res.* **1989**, *22*, 184.
- (25) Kollman, P. A. Free-energy calculations—Applications to chemical and biochemical phenomena. *Chem. Rev.* **1993**, *93*, 2395.
- (26) Hu, H.; Yang, W. T. Dual-topology/dual-coordinate free-energy simulation using QM/MM force field. *J. Chem. Phys.* **2005**, *123*, 041102.
- (27) Hu, H.; Yang, W. T. Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573.
- (28) Liu, S.; Pedersen, L. G. Estimation of Molecular Acidity via Electrostatic Potential at the Nucleus and Valence Natural Atomic Orbitals. *J. Phys. Chem. A* **2009**, *113*, 3648.
- (29) Giese, T. J.; York, D. M. Charge-dependent model for many-body polarization, exchange, and dispersion interactions in hybrid quantum mechanical/molecular mechanical calculations. *J. Chem. Phys.* **2007**, *127*, 194101.
- (30) Kastholz, M. A.; Huenenberger, P. H. Computation of methodology-independent ionic solvation free energies from molecular simulations. II. The hydration free energy of the sodium cation. *J. Chem. Phys.* **2006**, *124*, 224501.
- (31) Kastholz, M. A.; Hunenberger, P. H. Computation of methodology-independent ionic solvation free energies from molecular simulations. I. The electrostatic potential in molecular liquids. *J. Chem. Phys.* **2006**, *124*, 124106.
- (32) Reif, M. M.; Huenenberger, P. H. Computation of methodology-independent single-ion solvation properties from molecular simulations. III. Correction terms for the solvation free energies, enthalpies, entropies, heat capacities, volumes, compressibilities, and expansivities of solvated ions. *J. Chem. Phys.* **2011**, *134*, 144103.
- (33) Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420.

- (34) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Aqueous solvation free energies of ions and ion-water clusters based on an accurate value for the absolute aqueous solvation free energy of the proton. *J. Phys. Chem. B* **2006**, *110*, 16066.
- (35) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R. The proton's absolute aqueous enthalpy and Gibbs free energy of solvation from cluster-ion solvation data. *J. Phys. Chem. A* **1998**, *102*, 7787.
- (36) Hu, X. Q.; Hu, H.; Yang, W. T. *QM4D: An Integrated and Versatile Quantum Mechanical/Molecular Mechanical Simulation Package* (<http://www.qm4d.info/>, accessed 2012).
- (37) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648.
- (38) Lee, C.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti correlation energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785.
- (39) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926.
- (40) MacKerell, J., A.D.; Brooks, B.; Brooks, I., C.L.; Nilsson, L. R., B.; Won, Y.; Karplus, M., CHARMM: The energy function and its parameterization with an overview of the program. In *The Encyclopedia of Computational Chemistry*, Schleyer, P. v. R., Ed.; John Wiley & Sons: Chichester, U.K., 1998; Vol. 1, pp 271.
- (41) Hu, H.; Liu, H. Pitfall in QM/MM Molecular Dynamics Simulation of Small Solutes in Solution. *J. Phys. Chem. B* **2013**, *117*, 6505.
- (42) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684.
- (43) Zeng, X. C.; Hu, H.; Hu, X. Q.; Cohen, A. J.; Yang, W. T. Ab initio quantum mechanical/molecular mechanical simulation of electron transfer process: Fractional electron approach. *J. Chem. Phys.* **2008**, *128*, 124510.
- (44) Zeng, X. C.; Hu, H.; Hu, X. Q.; Yang, W. T. Calculating solution redox free energies with ab initio quantum mechanical/molecular mechanical minimum free energy path method. *J. Chem. Phys.* **2009**, *130*, 164111.
- (45) Wodrich, M. D.; Corminboeuf, C.; Schleyer, P. v. R. Systematic errors in computed alkane energies using B3LYP and other popular DFT functionals. *Org. Lett.* **2006**, *8*, 3631.
- (46) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. Insights into current limitations of density functional theory. *Science* **2008**, *321*, 792.
- (47) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112*, 289.
- (48) Riccardi, D.; Li, G.; Cui, Q. Importance of van der Waals Interactions in QM/MM Simulations. *J. Phys. Chem. B* **2004**, *108*, 6467.
- (49) Jin, Y.; Johnson, E. R.; Hu, X.; Yang, W.; Hu, H. Contributions of Pauli Repulsions to the Energetics and Physical Properties Computed in QM/MM Methods. Submitted to *J. Comput. Chem.*, DOI: 10.1002/jcc.23401, **2013**.
- (50) Cui, Q.; Karplus, M. Molecular properties from combined QM/MM methods. II. Chemical shifts in large molecules. *J. Phys. Chem. B* **2000**, *104*, 3721.
- (51) Brooks, C. L., Thermodynamic calculations in biological systems. In *Computer simulations of biomolecular systems*, van Gunsteren, W. F.; Weiner, P. K., Eds. ESCOM: Leiden, 1989; pp 73.
- (52) Wang, M. L.; Hu, X. Q.; Beratan, D. N.; Yang, W. T. Designing molecules by optimizing potentials. *J. Am. Chem. Soc.* **2006**, *128*, 3228.
- (53) Keinan, S.; Hu, X.; Beratan, D. N.; Yang, W. Designing molecules with optimal properties using the linear combination of atomic potentials approach in an AM1 semiempirical framework. *J. Phys. Chem. A* **2007**, *111*, 176.
- (54) Hu, X.; Beratan, D. N.; Yang, W. A gradient-directed Monte Carlo approach to molecular design. *J. Chem. Phys.* **2008**, *129*, 064102.
- (55) Keinan, S.; Therien, M. J.; Beratan, D. N.; Yang, W. Molecular Design of Porphyrin-Based Nonlinear Optical Materials. *J. Phys. Chem. A* **2008**, *112*, 12203.
- (56) Hu, X.; Beratan, D. N.; Yang, W. A gradient-directed Monte Carlo method for global optimization in a discrete space: Application to protein sequence design and folding. *J. Chem. Phys.* **2009**, *131*, 154117.
- (57) Hu, X.; Hu, H.; Beratan, D. N.; Yang, W. A Gradient-Directed Monte Carlo Approach for Protein Design. *J. Comput. Chem.* **2010**, *31*, 2164.
- (58) Jolly, W. L. *Modern Inorganic Chemistry*, 2nd Edition; McGraw-Hill: New York, 1991.
- (59) McMurry, J. E. *Fundamentals of Organic Chemistry*, 7th Edition; Cengage Learning: Independence, KY, and Victoria, Australia, 2010.