Calculating solution redox free energies with *ab initio* quantum mechanical/molecular mechanical minimum free energy path method

Xiancheng Zeng, Hao Hu, Xiangqian Hu, and Weitao Yang[a)]
Department of Chemistry, Duke University, Durham, North Carolina 27708, USA

(Received 12 November 2008; accepted 26 March 2009; published online 24 April 2009)

A quantum mechanical/molecular mechanical minimum free energy path (QM/MM-MFEP) method was developed to calculate the redox free energies of large systems in solution with greatly enhanced efficiency for conformation sampling. The QM/MM-MFEP method describes the thermodynamics of a system on the potential of mean force surface of the solute degrees of freedom. The molecular dynamics (MD) sampling is only carried out with the QM subsystem fixed. It thus avoids "on-the-fly" QM calculations and thus overcomes the high computational cost in the direct QM/MM MD sampling. In the applications to two metal complexes in aqueous solution, the new QM/MM-MFEP method yielded redox free energies in good agreement with those calculated from the direct QM/MM MD method. Two larger biologically important redox molecules, lumichrome and riboflavin, were further investigated to demonstrate the efficiency of the method. The enhanced efficiency and uncompromised accuracy are especially significant for biochemical systems. The QM/MM-MFEP method thus provides an efficient approach to free energy simulation of complex electron transfer reactions. © *2009 American Institute of Physics*.
[DOI: 10.1063/1.3120605]

I. INTRODUCTION

Calculating accurate free energies for many enzyme-catalyzed electron transfer (ET) processes is challenging because biological ET reactions often involve complicated interactions between enzyme, substrate, and cofactors.[1] The conformational dynamics makes calculating the redox potentials much more difficult than calculating the ionization potential in gas phase. Thus, many methods have been developed to compute the redox free energies. The continuum solvent model is one of the most popular approaches to describe the role of solvent in ET processes because of its simplicity.[2–10] Although the continuum solvent model can reveal essential physics of redox processes in solution, the dependence on parameters of the solute radii and the lack of the details of the structures or the thermodynamics motivated the employment of explicit solvent models in molecular dynamics (MD) simulation of redox reactions.[11–13] The pioneering works by Bader and co-workers[14,15] and Warshel and co-workers[16–20] used MD simulation to sample the explicit solvent phase space and provided insights of the free energy profiles for several solution ET processes. To reveal the structural and energetic changes of the redox molecules at the electronic level, Sprik and colleagues employed the Car–Parrinello MD method[21] to describe both the solute and the solvent and obtained accurate reorganization energies and geometrical information for solution ET reactions.[22–29] However, the high computational cost and finite system size limit the application of the *ab initio* MD method.[24,30]

With well balanced computational cost and accuracy, the combined quantum mechanical and molecular mechanical (QM/MM) method[31] becomes a very promising approach to simulate the biological ET processes, in which one treats the redox centers quantum mechanically and the rest of the system molecular mechanically.[32] In previous work, we have shown that the "on-the-fly" *ab initio* QM/MM calculation with direct MD sampling can be used to calculate the free energies for redox reactions in solution with satisfactory agreement with experimental data.[33] However, the cost of *ab initio* QM/MM calculation is still high, as such direct QM/MM MD simulations are only affordable with very limited timescales for small systems.[33–35] For large biomolecular systems, the computational cost for the QM calculation of the active site and phase space sampling of the protein severely limit the application of the direct QM/MM MD simulations. Although the semiempirical QM methods, including the empirical valence bond model[36] and the self-consistent-charge density functional tight-binding method,[37] can largely reduce the computational cost in the QM evaluations,[38–41] the reliability and transferability of the semiempirical QM methods still limited their applications.

The dual need for both accurate QM evaluations and broad phase space sampling is crucial for not only the ET reactions but also general reaction processes in solution and in enzymes. Therefore, several *ab initio* QM/MM simulation methods, including QM-FE,[42] frozen density functional theory (DFT),[43,44] and QM/MM-FE have been developed.[45] The QM/MM-FE method, developed in our laboratory, has shown advantages in many studies of enzymatic reactions.[35,46–53] However, the strong dependence of the QM geometry on the MM conformations limited its application to reactions in solution.[54,55] To eliminate this limitation, we recently developed the QM/MM minimum free energy path (QM/MM-MFEP) method,[55,56] which has been successfully applied to simulations of chemical reactions in solution and enzymes.[57,58] Through geometry optimizations of the QM

[a)]Electronic mail: yang@chem.duke.edu.

 © 2009 American Institute of Physics

subsystem on the potential of mean force (PMF) surface, the QM/MM-MFEP method generates a series of intermediate states with the minimum free energies along the reaction coordinate, thus avoids the expensive on-the-fly QM evaluations in the direct QM/MM MD sampling. It also provides a robust approach to obtaining well-defined reaction path in solution and in enzymes.

For redox processes, since we focus here on the thermodynamic free energy changes instead of the free energy barriers in chemical reactions, obtaining the entire MFEP is unnecessary, as the redox free energy is a state function and independent on the reaction path. Therefore, only the geometries of the reduced and the oxidized states need to be optimized on their corresponding PMF surfaces, then an arbitrary path can in principle be chosen to connect the end points for free energy calculation. With the optimized geometries of the end points, we can use the fractional number of electrons (FNE) to serve the reaction order parameter of the electronic states, which has been verified to be an effective approach to describe redox processes.[33] The linear interpolation of the geometries of the end points can be performed for the fractional electronic states to build a smooth alchemical reaction path for the free energy simulation of redox processes. In this manner, we even avoid the geometry optimization for the intermediate states and further reduce the computational cost of the free energy simulation.

In this article, we present the redox free energy simulation using the new QM/MM-MFEP method combined with the fractional electron approach to calculate the free energies for redox reactions in solution. This paper is organized as follows. We first briefly introduce the basic theory in the QM/MM-MFEP method. Then we show the practical procedures of the new method and the computational details of the simulation in Sec. III. In Sec. IV we compare the simulation results with our previous work and experiments, and also demonstrate the significantly enhanced efficiency in the QM/MM-MFEP method. Finally we conclude with Sec. V.

II. THEORY

A. Geometry optimization on the potential of mean force surface

To avoid performing the on-the-fly QM evaluations in the direct sampling of the phase space, we describe the thermodynamics of the solute with the PMF of the QM subsystem and only sample the phase space of the solvent with molecular mechanics.[55,56]

In the QM/MM representation, the PMF and the gradient of PMF with respect to the solute degrees of freedom can be written as

$$A(\mathbf{r}_{\mathrm{QM}}) = -\frac{1}{\beta}\ln\left[\int e^{-\beta E(\mathbf{r}_{\mathrm{QM}},\mathbf{r}_{\mathrm{MM}})} d\mathbf{r}_{\mathrm{MM}}\right], \tag{1}$$

$$\frac{\partial A(\mathbf{r}_{\mathrm{QM}})}{\partial \mathbf{r}_{\mathrm{QM}}} = \left\langle \frac{\partial E(\mathbf{r}_{\mathrm{QM}},\mathbf{r}_{\mathrm{MM}})}{\partial \mathbf{r}_{\mathrm{QM}}} \right\rangle_{E,\{\mathbf{r}_{\mathrm{MM}}\}}, \tag{2}$$

where $A(\mathbf{r}_{\mathrm{QM}})$ is the PMF, $\mathbf{r}_{\mathrm{QM}}$ is the solute geometry, $\mathbf{r}_{\mathrm{MM}}$ is the solvent conformation, $E(\mathbf{r}_{\mathrm{QM}},\mathbf{r}_{\mathrm{MM}})$ is the total potential energy of the molecular system, $1/\beta$ is the Boltzmann constant times temperature, and the symbol $\langle\cdots\rangle$ denotes the ensemble average.

Equations (1) and (2) build the foundation for geometry optimization on the PMF surface. Since the solute geometry and the solvent ensemble are mutually dependent, we can solve the self-consistent problem with iterative procedure. However, the straightforward implementation of the optimization method requires sampling of the solvent ensemble after every movement of QM geometry. To improve the computational efficiency, the optimization is instead performed in a fixed ensemble.[55,56] The first snapshot of the QM geometry is set as a reference, and the free energy is defined to be relative to the reference

$$A(\mathbf{r}_{\mathrm{QM}}) = A_{\mathrm{ref}} - \frac{1}{\beta}\ln\langle e^{-\beta[E(\mathbf{r}_{\mathrm{QM}},\mathbf{r}_{\mathrm{MM}})-E_{\mathrm{ref}}]}\rangle_{E_{\mathrm{ref}},\{\mathbf{r}_{\mathrm{MM}}\}}, \tag{3}$$

where E_{ref} is the potential energy of the reference geometry, i.e., $E(\mathbf{r}_{\mathrm{QM}}^{\mathrm{ref}},\mathbf{r}_{\mathrm{MM}})$, A_{ref} is the free energy of the reference, and $\{\mathbf{r}_{\mathrm{MM}}\}$ is the ensemble sampled with the reference QM geometry. According to Eq. (3), the PMF gradient in this method is then[55]

$$\frac{\partial A(\mathbf{r}_{\mathrm{QM}})}{\partial \mathbf{r}_{\mathrm{QM}}} = \frac{\left\langle \dfrac{\partial E(\mathbf{r}_{\mathrm{QM}},\mathbf{r}_{\mathrm{MM}})}{\partial \mathbf{r}_{\mathrm{QM}}} e^{-\beta[E(\mathbf{r}_{\mathrm{QM}},\mathbf{r}_{\mathrm{MM}})-E_{\mathrm{ref}}]} \right\rangle_{E_{\mathrm{ref}},\{\mathbf{r}_{\mathrm{MM}}\}}}{\langle e^{-\beta[E(\mathbf{r}_{\mathrm{QM}},\mathbf{r}_{\mathrm{MM}})-E_{\mathrm{ref}}]}\rangle_{E_{\mathrm{ref}},\{\mathbf{r}_{\mathrm{MM}}\}}}. \tag{4}$$

The PMF gradient in Eq. (4) gives the steepest direction to the minimum. When the distribution of $e^{-\beta[E(\mathbf{r}_{\mathrm{QM}})-E_{\mathrm{ref}}]}$ is converged, one can use various optimization methods to minimize the free energy. The optimized QM geometry is then used to generate a new ensemble of MM conformations in next cycle. This procedure can be iterated until convergence is reached.

Utilization of a finite reference ensemble, i.e., Eqs. (3) and (4), requires special technical considerations in practice. Since the free energy is computed by the free energy perturbation (FEP) method [Eq. (3)], the accuracy of ΔA depends on the overlapping of the phase space of two states. If the phase spaces of two states differ significantly, the computed ΔA cannot be trusted. Similar problems have been in fact discussed in the case of Jarzynski fast-growth simulation method.[59–61] Because of this problem, in FEP simulations with classical force fields, an empirical rule is often implemented such that ΔA of a single FEP simulation cannot be larger than $2kT$;[60] however, changes in electronic state without much change of geometry can easily cause energy fluctuations more than a few kT. According to cumulant expansion, the free energy can be expanded to the second order as

$$\Delta A = \langle E \rangle - \frac{\beta}{2}(\langle E^2 \rangle - \langle E \rangle^2) + \mathcal{O}(E^3), \tag{5}$$

$$= \langle E \rangle - \frac{\beta}{2}\sigma^2 + \mathcal{O}(E^3), \tag{6}$$

where $\sigma^2 = \langle E^2 \rangle - \langle E \rangle^2$ is the fluctuation of potential energies. When we minimize the free energy ΔA, the average potential $\langle E \rangle$ is minimized while the fluctuation σ^2 is maximized. For systems with flat potential surfaces, the QM geometries will be lead to conformations with maximum statistic errors. We here applied another approach to resolve this problem. In all geometry optimization algorithms currently employed in our study and implemented in popular packages such as GAUSSIAN03, the goal is to identify the local minimum within a single state. In such a case, one can make a valid assumption that the entropy of any point in this state is constant. Therefore, one may minimizes ΔH, instead of ΔA, in our QM/MM-MFEP optimizations. Under this assumption, the objection function for geometry minimization is $\langle E(\mathbf{r}_{\mathrm{QM}}, \mathbf{r}_{\mathrm{MM}}) \rangle$ with the gradient

$$\frac{\partial \langle E(\mathbf{r}_{\mathrm{QM}}, \mathbf{r}_{\mathrm{MM}}) \rangle}{\partial r_{\mathrm{QM}}} = \frac{\int \frac{\partial E(\mathbf{r}_{\mathrm{QM}}, \mathbf{r}_{\mathrm{MM}})}{\partial \mathbf{r}_{\mathrm{QM}}} e^{-\beta E_{\mathrm{ref}}} d\mathbf{r}_{\mathrm{MM}}}{\int e^{-\beta E_{\mathrm{ref}}} d\mathbf{r}_{\mathrm{MM}}}. \tag{7}$$

In this manner, the higher order terms are excluded in the minimization and the statistic errors are thus reduced.

B. Redox free energy calculation using fractional electron approach

Calculating accurate free energy difference between two distinct states usually requires some techniques such as FEP or thermodynamic integration (TI) methods. To characterize the electronic states, most redox free energy simulations introduced the mixing coefficient of the energy functions of the two states,[20,22,39,40,62] or the FNE as the reaction order parameter.[33,63–65] Because the free energy is a state function, in principle, the choice of the order parameter will not change the free energy difference. However, different choices of the order parameter will affect the convergence of the simulations and accuracy of the final results.[66]

Our previous work has implemented the FNE as a novel order parameter for the QM/MM TI calculation of redox free energies. In this work, we also use the FNE to connect the reduced and the oxidized states of the solute and use the FEP method to calculate accurate redox free energy. Specifically, we here choose the order parameter as a combination of the FNE and the geometrical difference between the two final states. That is, after the structural optimization for the two end states with the QM/MM-MFEP method, the intermediate structures were generated by linear interpolation between the structures of the two end states; each intermediate state was also assigned a FNE linearly interpolated from the changes of electron number in the redox processes. Although one could use only the FNE to define the intermediate states and correspondingly optimize the structures for each state before the FEP simulation. Our approach will be computationally more efficient because of the smooth structural change between different intermediate states and the elimination of the geometry optimization on the QM PMF for the intermediate states.

In the FEP method, we then divide the whole ET process into several intervals and perform sampling on a series of the intermediate states $\{\lambda_i\}_{i=0,\ldots,n}$, characterized by different values of the FNE and the corresponding geometries obtained from the linear interpolation. The overall free energy difference can be obtained by summing up the free energy change in each sampling interval,

$$\Delta A = \sum_{i=0}^{n-1} \Delta A(\lambda_i \rightarrow \lambda_{i+1}), \tag{8}$$

$$= \sum_{i=0}^{n-1} -\frac{1}{\beta} \ln[\langle e^{-\beta \Delta E(\lambda_i \rightarrow \lambda_{i+1})} \rangle_{\lambda_i}], \tag{9}$$

where λ is the reaction order parameter, i denotes the ith intermediate state, n is the total number of states sampled, and the symbol $\langle \cdots \rangle_{\lambda_i}$ designates the ith ensemble. $i=0$ denotes the reactant state, and $i=n$ denotes the product state. $\Delta E(\lambda_i \rightarrow \lambda_{i+1})$ is the potential energy difference between two electronic states, which can be computed with the QM/MM-FE approach developed in previous work.[45,55,56] By turning λ_0 to λ_n, we drive the system from the reduced electronic state to the oxidized state. Although the ΔA obtained in Eq. (9) does not include the zero-point energy and the vibrational contribution to the redox free energy, those terms can be estimated using frequency analysis within the harmonic approximation.

III. COMPUTATIONAL DETAILS

To validate the QM/MM-MFEP method for the calculation of redox free energies, we first applied the method to simulate the aqueous iron and ruthenium complexes, which have been studied using the more substantiated direct QM/MM MD method in our previous work.[33] For further demonstration of the efficiency and accuracy of the QM/MM-MFEP method, we performed two additional simulations for flavin derivatives: lumichrome (LC) and riboflavin (RF). The entire LC and RF molecules are treated quantum mechanically and their chemical structures are shown in Fig. 1.

The geometries of the solute were optimized for both the reduced and the oxidized states on their corresponding PMF surfaces. The iterative sequential sampling and optimization followed the procedures in our previous paper.[55] In the iterative geometry optimizations, each MD sampling was performed for 64 ps after a 16 ps equilibration.

After the solute geometries were optimized for both reduced and oxidized states, the redox process was divided into 11 sampling intervals for the FEP calculation, i.e., $n=11$ in Eq. (8), and the linear interpolation of the solute geometries was performed to generate a series of geometries of the QM subsystem with fractional electrons to smoothly connect the end points. To calculate the free energy, a new MD sampling was performed for each intermediate state

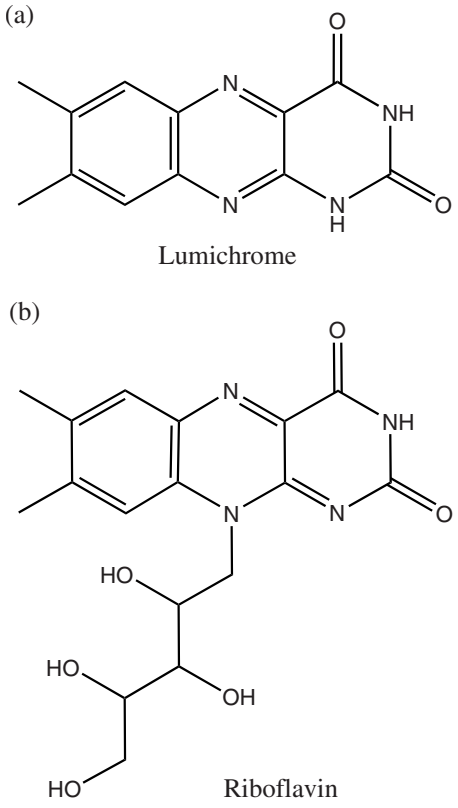

FIG. 1. Molecular diagrams of flavin derivatives: (a) LC and (b) RF. The LC and RF molecules are treated quantum mechanically in the simulation.

$r_{\mathrm{QM},\lambda_i}$ with the solute in the corresponding FNE state. In the FEP calculation, the MD sampling was increased to 128 ps after a 32 ps equilibration for better statistic convergence. Finally, the free energy contributions from zero-point energy and thermodynamic vibration were estimated by frequency analysis using the GAUSSIAN03 program.[67]

The MD simulations were carried out with the program SIGMA,[68,69] while the QM calculations were performed with GAUSSIAN03.[67] The electrostatic potential fit charges of QM subsystem were obtained with our recently developed method,[70] which improves the numerical stability of the charges on the atoms in the QM region. Since the Ewald summation for charged systems with a local basis set has not been implemented, we here use a QM-MM interaction distance cutoff. Following the protocol in the previous work,[55] the dual cutoffs were used in the simulations, with the short cutoff to be 14 Å and the long cutoff to be 28 and 20 Å for the metal ions and organic molecules, respectively. The long cutoff was chosen as such to ensure the convergence of QM/MM electrostatic interactions and the final redox free energies.[33,71,72] The cutoff for the organic systems were chosen as 20 Å because the net charge of the system is between 0 and −1, which is smaller than that of the metal systems. Additional FEP calculation with 25 Å cutoff for the LC molecule was performed to demonstrate the convergence of the long distance interactions (see Sec. IV).

The metal complexes were modeled with the same configuration in the previous work:[33] The ions and six coordinating water molecules were treated quantum mechanically and were solvated in cubic boxes with the dimensions $64\times64\times64$ Å^3, which contained 8.6×10^3 TIP3P water molecules as the solvent.[12] For the flavin derivatives, the LC and RF molecules are treated quantum mechanically and solvated in cubic boxes with the dimensions: $64\times64\times64$ Å^3 for LC and $54\times62\times62$ Å^3 for RF, respectively. The dimensions of the boxes were chosen such that the solutes were surrounded by solvent layers with thickness at least greater than the corresponding long cutoff. All systems were simulated under *NVT* conditions with temperature $T=300$ K.

IV. RESULTS AND DISCUSSIONS

A. Geometry optimization on the PMF

In the QM/MM-MFEP method, each geometry optimization was performed on the PMF surface in a fixed MM ensemble. The MM ensemble was then updated with the optimized solute geometry. The converged MM ensemble and the corresponding fully optimized QM geometry is critical to the accuracy and reliability of the method. Figure 2 shows the convergence of the four systems in the oxidized states in terms of the free energy variation. The horizontal axis is the number of QM evaluations in the geometry optimization. Each dashed line denotes the termination of one QM optimization cycle, where a new MM ensemble was sampled. The free energies of the systems converged to plateaus after 8–10 cycles of iterative sequential sampling and geometry optimization for both the aqueous metal complexes and the flavin derivatives. The spikes in the free energy variations were generated by the optimization algorithm employed in the GAUSSIAN03 program, which are frequently observed in various systems.[55]

B. Redox free energies

With the QM subsystems optimized on the PMF surfaces in fully converged MM ensembles, the FEP simulation was then performed to evaluate the redox free energies for the systems. The oxidation free energies of $Fe(H_2O)_6^{2/3+}$ and $Ru(H_2O)_6^{2/3+}$ obtained from Eq. (9) are 5.70 and 5.01 eV, respectively. These values do not include the contributions from vibrational entropy and zero-point energies.

With the harmonic approximation, frequency calculations can be performed in the QM/MM-MFEP method to estimate the free energy contribution from the solute vibrational entropy ΔG_{vib}. In the direct QM/MM MD method, the vibrational dynamics of QM subsystem is represented by the classical motions of solute from MD sampling. Thus, the vibrational contribution is already included. In the QM/MM-MFEP method, we estimated the vibrational contributions of the solutes through the frequency calculations on the enthalpy surfaces. The vibrational frequencies of the QM subsystems were calculated in the mean field of the solvent that is represented by MM point charges. The ΔG_{vib} for $Fe(H_2O)_6^{2/3+}$ and $Ru(H_2O)_6^{2/3+}$ are estimated to be 0.042 and 0.027 eV, respectively. After the correction, the oxidation free energies are 5.74 eV for $Fe(H_2O)_6^{2/3+}$ and 5.04 eV for $Ru(H_2O)_6^{2/3+}$. As shown in Table I, these values are comparable to the results obtained in a previous study using the direct QM/MM MD method, where the oxidation free energies were calculated to be 5.82 eV for $Fe(H_2O)_6^{2/3+}$ and 5.14

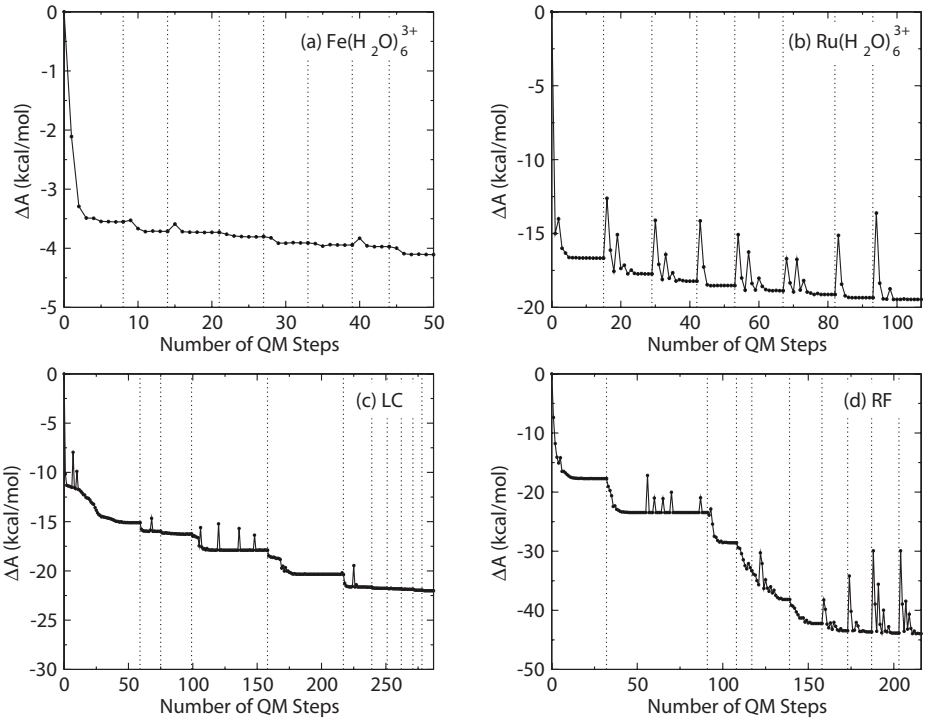

FIG. 2. Convergence of the relative free energies during the geometry optimization for (a) $Fe(H_2O)_6^{3+}$, (b) $Ru(H_2O)_6^{3+}$, (c) LC, and (d) RF in QM/MM-MFEP method. The reduced states of the systems converge similarly as the oxidized states shown here. The vertical dash lines separate different MM ensembles obtained by sequential sampling. Within the region between two dash lines, the dotted lines are the relative free energies of the systems during the QM optimizations on their PMF surfaces in the corresponding MM ensembles. The systems converge after 8–10 cycles of MM sampling and QM optimizations.

eV for $Ru(H_2O)_6^{2/3+}$. The discrepancy between the two methods is 0.07–0.10 eV or 1.6–2.3 kcal/mol, which is within the statistic deviations.

Besides the vibrational entropy, another missing term is the difference in the zero-point energy between the two redox states. Since the zero-point energy calculation in the direct QM/MM MD method is expensive but not dominant, this term was omitted. However, in the QM/MM-MFEP method, the zero-point energy of different redox states can be easily estimated through frequency calculations on the optimized geometries. The free energy corrections from zero-point energy ΔG_{ZPE} are about 0.10 eV and 0.06 eV for $Fe(H_2O)_6^{2/3+}$ and $Ru(H_2O)_6^{2/3+}$, respectively. Including the correction from both vibrational entropy and zero-point energy, the absolute oxidation free energies ΔG_{abs} obtained from the QM/MM-MFEP method for $Fe(H_2O)_6^{2/3+}$ and $Ru(H_2O)_6^{2/3+}$ are 5.84 and 5.10 eV, respectively.

The experimental values are measured relative to the standard hydrogen electrode (SHE), which is artificially set to be zero. Therefore, the absolute value of the oxidation free energies for SHE (ΔG_{abs}^{SHE}) should be added in the experimental data when we compare the theoretical values to experiments. However, the measurements of ΔG_{abs}^{SHE} are not very accurate and disputable, with a range from 4.2 ± 0.4 to 4.84 eV.[73–77] In Tables I and II, the experimental values of the absolute oxidation free energies are listed as two sets of data with the minimum and maximum values of ΔG_{abs}^{SHE}.

Because the computational cost in the simulation of LC and RF molecules using the direct QM/MM MD method is very expensive, the two bio-organic systems were only simulated using the QM/MM-MFEP method. As shown in Table II, the oxidation free energies of the redox couples $LC/LC^{\cdot-}$ and $RF/RF^{\cdot-}$ are evaluated to be 3.16 and 3.41 eV, respectively. Since we here used a shorter cutoff (20 Å) than that of metal systems (28 Å), we performed FEP calculations with 25 Å cutoff for $LC/LC^{\cdot-}$. The oxidation free energy is 3.12 eV with 25 Å cutoff and 3.16 eV with 20 Å cutoff, respectively. The 0.04 eV difference between the FEP calculations suggests that the long range electrostatic potential converged at the distance of 20–25 Å. The correction from ΔG_{vib} and

TABLE I. Oxidation free energies of metal complexes calculated with direct QM/MM MD and QM/MM-MFEP methods. Units are in eV.

		ΔG	ΔG_{vib}	$\Delta G+\Delta G_{vib}$	ΔG_{ZPE}	ΔG_{abs}	ΔG_{abs}^{expt}
$Fe(H_2O)_6^{2/3+}$	Direct	···	···	5.82[a]	···	···	
	MFEP	5.70	0.04	5.74[b]	0.10	5.84[c]	5.0/5.6[d]
$Ru(H_2O)_6^{2/3+}$	Direct	···	···	5.14[a]	···	···	
	MFEP	5.01	0.03	5.04[b]	0.06	5.10[c]	4.4/5.1[d]

[a]Values reported in Ref. 33, which already included the free energy contributions from vibrational dynamics.
[b]In order to compare with the results using the direct QM/MM MD method, the vibrational contributions ΔG_{vib} are estimated and added to the results for QM/MM-MFEP method.
[c]ΔG_{abs} are the absolute redox free energies, included the corrections from both vibrational dynamics ΔG_{vib} and zero-point energies ΔG_{ZPE}. The results using direct QM/MM MD method did not include ΔG_{ZPE}.
[d]The experimental absolute free energies are obtained by taking absolute oxidation free energy for SHE to be from 4.2 to 4.84 eV (Refs. 73–77).

TABLE II. Comparison of oxidation free energies of LC and RF calculated with QM/MM-MFEP methods to experiments. Units are in eV.

	MFEP[a]			Experimental data[b]		$\Delta\Delta G$ [c]	
	ΔG	$\Delta G_{\text{vib+ZPE}}$	ΔG_{abs}	ΔG^0	$\Delta G_{\text{abs}}^{\text{expt}}$	MFEP	Expt.
$LC/LC^{\cdot-}$	3.16	0.07	3.23	−0.502	3.7/4.3	···	···
$RF/RF^{\cdot-}$	3.41	0.07	3.48	−0.292	3.9/4.5	0.25	0.21

[a]ΔG are the values before any correction. $\Delta G_{\text{vib+ZPE}}$ are the corrections considering the vibrational contributions and zero-point energies. ΔG_{abs} are values of absolute oxidation free energies including the corrections from both vibrational contributions and zero-point energies.
[b]Experimental data are based on the standard redox free energies ΔG^0 from Ref. 82, which are relative to the SHE. $\Delta G_{\text{abs}}^{\text{SHE}}$, the values of the absolute oxidation free energy for SHE, were measured to be from 4.2 to 4.84 eV (Refs. 73–77). The experimental absolute free energies are therefore obtained as $\Delta G_{\text{abs}}^{\text{expt}}=\Delta G^0+\Delta G_{\text{abs}}^{\text{SHE}}$.
[c]Relative oxidation free energy difference between the LC and RF molecules.

ΔG_{ZPE} is coincidentally to be about 0.07 eV for both LC and RF systems, which leads to 3.23 and 3.48 eV as the absolute oxidation free energies for LC and RF, respectively.

The experimental oxidation potentials of the $LC/LC^{\cdot-}$ and $RF/RF^{\cdot-}$ couples are −0.502 and −0.292 V. Taking the Nernst equation ($\Delta G=-nF\Delta E$), the relative oxidation free energy difference $\Delta\Delta G$ between the two couples is 0.21 eV. The result obtained from QM/MM-MFEP method is 0.25 eV, which is in excellent accordance with the experimental value.

C. Computational efficiency

The agreement between the oxidation free energies calculated using the two methods indicates that the QM/MM-MFEP method is at least as accurate as the direct QM/MM MD method. In the Table III, we compared the efficiency between the direct QM/MM MD and the QM/MM-MFEP methods when performing similar simulations to calculate redox free energies. Using LanL2DZ basis set,[78] the aqueous metal complexes contain 100 basis functions. The direct QM/MM MD simulation spent about 40 days on dual Intel Xeon 3.60 GHz CPUs to complete 20 ps MD sampling with 20 000 times of the on-the-fly QM evaluations. The QM/MM-MFEP simulations on the same systems performed about 8 cycles of QM optimization and sequential MD sampling in 20 days, with each cycle consisting of 80 ps sampling. If we want to perform the direct QM/MM MD simulations for 20 ps on LC or RF molecule, the estimated computational time would exceed 400 days. In contrast, the QM/MM-MFEP simulation of the LC and RF molecules were completed in 25 and 45 days, respectively. As the system becomes larger, the advantage of the QM/MM-MFEP method is more significant. The computational efficiency is promoted about 15–30 folds for the two bio-organic systems because the on-the-fly QM evaluations are avoided and the thermodynamics is described by the PMF of the QM subsystem. For the systems in this paper, the linear interpolation was performed to build the path for free energy simulation. However, we need to point out that a smooth connection between the redox states through the FNE cannot be guaranteed in all cases. Caution may be needed in the simulation of special systems with FNE to reduce possible artifacts caused by DFT method.[30,79–81] Also, if the interpolated reaction path is bumpy, e.g., for reactions with significant geometrical changes between the redox states, the optimization for a smooth path is desired to reach the convergence. Even if the optimization of a reaction path rather than an interpolated path is necessary, the computational cost in terms of time may not significantly increase since the optimization of the intermediate states can be performed parallelly.

TABLE III. Efficiency comparison between direct QM/MM MD and QM/MM-MFEP methods.

Systems[a]	$M(H_2O)_6^{2/3+}$		$LC/LC^{\cdot-}$		$RF/RF^{\cdot-}$	
Total atoms	19		28		47	
Basis set	LanL2DZ		6-31+G**		6-31+G**	
Number of basis	100		392		613	
QM evaluation time[b]	3 min		30 min		100 min	
Method	Direct	MFEP	Direct[c]	MFEP	Direct[c]	MFEP
Sampling length	20 ps	80 ps	20 ps	80 ps	20 ps	80 ps
QM evaluation number	20 000×6	100×2 [d]	20 000×6 [e]	300×2 [d]	20 000×6 [e]	300×2 [d]
Total real time[f]	40 days	20 days	400 days	25 days	1400 days	45 days

[a]M=Fe,Ru.
[b]These are the approximate time costs for a single QM evaluation. The time costs here are based on running GAUSSIAN03 program on dual Intel Xeon 3.60 GHz CPUs.
[c]The direct QM/MM MD simulations are not performed for LC and RF. The corresponding parameters and time costs are estimations only.
[d]The geometry optimizations are only necessary for the reduced and the oxidized state of solute.
[e]Assume using the same setup of simulations as in Ref. 33, where TI method with six sampling intervals is used.
[f]These are the approximate time costs for performing the entire simulations.

V. CONCLUSION

In this work, we have implemented the QM/MM-MFEP method to calculate redox free energies, with significantly enhanced efficiency and uncompromised accuracy. We have verified the validity and accuracy of the QM/MM-MFEP method by repeating the calculation of redox free energies for $Fe(H_2O)_6^{2/3+}$ and $Ru(H_2O)_6^{2/3+}$ aqueous complexes. The results from the QM/MM-MFEP method agree well with those from the direct QM/MM MD method, with about 0.1 eV difference in the oxidation free energy. To demonstrate the efficiency of the QM/MM-MFEP method, two larger biochemical molecules, LC and RF, are further investigated. The oxidation free energies of the redox couples $LC/LC^{\cdot-}$ and $RF/RF^{\cdot-}$ were evaluated to be 3.23 and 3.48 eV, respectively. The relative free energy difference $\Delta\Delta G$ is 0.25 eV, which is very close to the experimental value of 0.21 eV. The computational time for the two bio-organic systems is shortened by about 15–30 folds compared to the direct QM/MM MD. The high accuracy and the enhanced MD sampling of the QM/MM-MFEP method shows promise for applications of the studies on redox reactions in biochemistry.

ACKNOWLEDGMENTS

Support from the National Institutes of Health, Project No. 5R01GM061870-08, is gratefully appreciated.

[1] M. Bixon and J. Jortner, Adv. Chem. Phys. **106**, 1 (1999).
[2] R. A. Marcus, J. Chem. Phys. **24**, 966 (1956).
[3] R. A. Marcus, J. Chem. Phys. **24**, 979 (1956).
[4] M. Z. Born, Physik **1**, 45 (1920).
[5] L. Onsager, J. Am. Chem. Soc. **58**, 1486 (1936).
[6] J. G. Kirkwood, J. Chem. Phys. **7**, 911 (1939).
[7] K. Rosso and J. Rustad, J. Phys. Chem. A **104**, 6718 (2000).
[8] M. Uudsemaa and T. Tamm, J. Phys. Chem. **107**, 9997 (2003).
[9] P. Jaque, A. Marenich, C. Cramer, and D. Truhlar, J. Phys. Chem. C **111**, 5783 (2007).
[10] A. Warshel, A. Papazyan, and I. Muegge, JBIC, J. Biol. Inorg. Chem. **2**, 143 (1997).
[11] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, *Intermolecular Forces* (Reidel, Dordrecht, Holland, 1981).
[12] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[13] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
[14] R. A. Kuharski, J. S. Bader, D. Chandler, M. Sprik, M. L. Klein, and R. W. Impey, J. Chem. Phys. **89**, 3248 (1988).
[15] J. S. Bader and D. Chandler, J. Phys. Chem. **96**, 6423 (1992).
[16] A. Warshel, J. Phys. Chem. **86**, 2218 (1982).
[17] A. Warshel and J.-K. Hwang, J. Chem. Phys. **84**, 4938 (1986).
[18] J.-K. Hwang and A. Warshel, J. Am. Chem. Soc. **109**, 715 (1987).
[19] J. K. Hwang, G. King, S. Greighton, and A. Warshel, J. Am. Chem. Soc. **110**, 5297 (1988).
[20] G. King and A. Warshel, J. Chem. Phys. **93**, 8682 (1990).
[21] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[22] J. Blumberger and M. Sprik, J. Phys. Chem. B **108**, 6529 (2004).
[23] J. Blumberger, L. Bernasconi, I. Tavernelli, R. Vuilleumier, and M. Sprik, J. Am. Chem. Soc. **126**, 3928 (2004).
[24] J. Blumberger and M. Sprik, J. Phys. Chem. B **109**, 6793 (2005).
[25] Y. Tateyama, J. Blumberger, M. Sprik, and I. Tavernelli, J. Chem. Phys. **122**, 234505 (2005).
[26] J. Blumberger, I. Tavernelli, M. L. Klein, and M. Sprik, J. Chem. Phys. **124**, 064507 (2006).
[27] M. Sulpizi, S. Raugei, J. VandeVondele, P. Carloni, and M. Sprik, J. Phys. Chem. B **111**, 3969 (2007).
[28] F. Costanzo, M. Sulpizi, R. Guido Della Valle, and M. Sprik, J. Chem. Theory Comput. **4**, 1049 (2008).
[29] J. Blumberger, J. Am. Chem. Soc. **130**, 16065 (2008).
[30] R. Ayala and M. Sprik, J. Phys. Chem. B **112**, 257 (2008).
[31] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).
[32] J. Blumberger, Phys. Chem. Chem. Phys. **10**, 5651 (2008).
[33] X. Zeng, H. Hu, X. Hu, A. J. Cohen, and W. Yang, J. Chem. Phys. **128**, 124510 (2008).
[34] M. Cascella, A. Magistrato, I. Tavernelli, P. Carloni, and U. Rothlisberger, Proc. Natl. Acad. Sci. U.S.A. **103**, 19641 (2006).
[35] S. Wang, P. Hu, and Y. Zhang, J. Phys. Chem. B **111**, 3758 (2007).
[36] A. Warshel and R. M. Weiss, J. Am. Chem. Soc. **102**, 6218 (1980).
[37] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, Phys. Rev. B **58**, 7260 (1998).
[38] M. H. Olsson, P. K. Sharma, and A. Warshel, FEBS Lett. **579**, 2026 (2005).
[39] M. S. Formaneck, G. Li, X. Zhang, and Q. Cui, J. Theor. Comput. Chem. **1**, 53 (2002).
[40] G. Li, X. Zhang, and Q. Cui, J. Phys. Chem. B **107**, 8643 (2003).
[41] S. Bhattacharyya, M. Stankovich, D. Truhlar, and J. Gao, J. Phys. Chem. A **111**, 5729 (2007).
[42] J. Chandrasekhar, S. F. Smith, and W. L. Jorgensen, J. Am. Chem. Soc. **107**, 154 (1985).
[43] T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).
[44] M. Olsson, G. Hong, and A. Warshel, J. Am. Chem. Soc. **125**, 5025 (2003).
[45] Y. Zhang, H. Liu, and W. Yang, J. Chem. Phys. **112**, 3483 (2000).
[46] G. Cisneros, H. Liu, Y. Zhang, and W. Yang, J. Am. Chem. Soc. **125**, 10384 (2003).
[47] P. Hu and Y. Zhang, J. Am. Chem. Soc. **128**, 1272 (2006).
[48] L. Wang, X. Yu, P. Hu, S. Broyde, and Y. Zhang, J. Am. Chem. Soc. **129**, 4731 (2007).
[49] M. Wang, Z. Lu, and W. Yang, J. Chem. Phys. **121**, 101 (2004).
[50] G. Cisneros, M. Wang, P. Silinski, M. Fitzgerald, and W. Yang, Biochemistry **43**, 6885 (2004).
[51] M. Wang, Z. Lu, and W. Yang, J. Chem. Phys. **124**, 124516 (2006).
[52] C. Corminboeuf, P. Hu, M. Tuckerman, and Y. Zhang, J. Am. Chem. Soc. **128**, 4530 (2006).
[53] J. Kastner, H. Senn, S. Thiel, N. Otte, and W. Thiel, J. Chem. Theory Comput. **2**, 452 (2006).
[54] M. Klahn, S. Braun-Sand, E. Rosta, and A. Warshel, J. Phys. Chem. B **109**, 15645 (2005).
[55] H. Hu, Z. Lu, J. M. Parks, S. K. Burger, and W. Yang, J. Chem. Phys. **128**, 034105 (2008).
[56] H. Hu, Z. Lu, and W. Yang, J. Chem. Theory Comput. **3**, 390 (2007).
[57] H. Hu, A. Boone, and W. Yang, J. Am. Chem. Soc. **130**, 14493 (2008).
[58] J. M. Parks, H. Hu, J. Rudolph, and W. Yang, J. Phys. Chem. B (submitted).
[59] G. Hummer and A. Szabo, Proc. Natl. Acad. Sci. U.S.A. **98**, 3658 (2001).
[60] H. Hu, R.-H. Yun, and J. Hermans, Mol. Simul. **28**, 67 (2002).
[61] S. Park and K. Schulten, J. Chem. Phys. **120**, 5946 (2004).
[62] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).
[63] R. Vuilleumier, M. Sprik, and A. Alavi, J. Mol. Struct. **506**, 343 (2000).
[64] I. Tavernelli, R. Vuilleumier, and M. Sprik, Phys. Rev. Lett. **88**, 213002 (2002).
[65] P. H.-L. Sit, M. Cococcioni, and N. Marzari, Phys. Rev. Lett. **97**, 028303 (2006).
[66] C. Dellago, P. G. Bolhuis, and D. Chandler, J. Chem. Phys. **108**, 9236 (1998).
[67] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision D.02, Gaussian, Inc., Wallingford, CT, 2004.
[68] H. Hu, M. Elstner, and J. Hermans, Proteins **50**, 451 (2003).
[69] G. Mann, R. H. Yun, L. Nyland, J. Prins, J. Board, and J. Hermans, *Proceedings of the 3rd International Workshop on Algorithms for Macromolecular Modelling* (Springer-Verlag, Berlin, 2002).
[70] H. Hu, Z. Lu, and W. Yang, J. Chem. Theory Comput. **3**, 1004 (2007).
[71] G. Hummer, L. R. Pratt, and A. E. Garcia, J. Phys. Chem. A **102**, 7885 (1998).
[72] See EPAPS Document No. E-JCPSA6-130-004917 for the phase potential effect of the cutoff treatment. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[73] W. Donald, R. Leib, J. O'Brien, M. Bush, and E. Williams, J. Am. Chem. Soc. **130**, 3371 (2008).

[74] C. Kelly, C. Cramer, and D. Truhlar, J. Phys. Chem. B **110**, 16066 (2006).
[75] H. Reiss, J. Phys. Chem. **89**, 4207 (1985).
[76] R. Gomer, J. Chem. Phys. **66**, 4413 (1977).
[77] Y. Marcus, J. Chem. Soc., Faraday Trans. **87**, 2995 (1991).
[78] P. J. Hay and W. R. Wadt, J. Chem. Phys. **82**, 270 (1985).
[79] J. VandeVondele and M. Sprik, Phys. Chem. Chem. Phys. **7**, 1363 (2005).
[80] P. Mori-Sanchez, A. J. Cohen, and W. Yang, Phys. Rev. Lett. **100**, 146401 (2008).
[81] A. J. Cohen, P. Mori-Sanchez, and W. Yang, Science **321**, 792 (2008).
[82] P. Wardman, J. Phys. Chem. Ref. Data **18**, 1637 (1989).