Accelerating QM/MM Free Energy Calculations: Representing the Surroundings by an Updated Mean Charge Distribution

Edina Rosta,[†,‡] Maciej Haranczyk,[†,§] Zhen T. Chu,[†] and Arieh Warshel*[,†]

Department of Chemistry, University of Southern California, 418 SGM Building, 3620 McClintock Avenue, Los Angeles, California 90089-1062, Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (NIH), Bethesda, Maryland 20892-0520, and Department of Chemistry, University of Gdańsk, 80-952 Gdańsk, Poland

Received: December 6, 2007; In Final Form: January 24, 2008

Reliable studies of enzymatic reactions by combined quantum mechanical/molecular mechanics (QM(ai)/MM) approaches with an ab initio description of the quantum region presents a major challenge to computational chemists. The main problem is the need for very large computer time to evaluate the QM energy, which in turn makes it extremely challenging to perform proper configurational sampling. One of the most obvious options for accelerating QM/MM simulations is the use of an average solvent potential. In fact, the idea of using an average solvent potential is rather obvious and has implicitly been used in Langevin dipole/QM calculations. However, in the case of explicit solvent models the practical implementations are more challenging, and the accuracy of the averaging approach has not been validated. The present study introduces the average effect of the fluctuating solvent charges by using equivalent charge distributions, which are updated every *m* steps. Several models are evaluated in terms of the resulting accuracy and efficiency. The most effective model divides the system into an inner region with *N* explicit solvent atoms and an external region with two effective charges. Different models are considered in terms of the division of the solvent system and the update frequency. Another key element of our approach is the use of the free energy perturbation (FEP) and/or linear response approximation treatments that guarantees the evaluation of the rigorous solvation free energy. Special attention is paid to the convergence of the calculated solvation free energies and the corresponding solute polarization. The performance of the method is examined by evaluating the solvation of a water molecule and a formate ion in water and also the dipole moment of water in water solution. Remarkably, it is found that different averaging procedures eventually converge to the same value but some protocols provide optimal ways of obtaining the final QM(ai)/MM converged results. The current method can provide computational time saving of 1000 for properly converging simulations relative to calculations that evaluate the QM(ai)/MM energy every time step. A specialized version of our approach that starts with a classical FEP charging and then evaluates the free energy of moving from the classical potential to the QM/MM potential appears to be particularly effective. This approach should provide a very powerful tool for QM-(ai)/MM evaluation of solvation free energies in aqueous solutions and proteins.

I. Introduction

Quantum mechanical/molecular mechanics (QM/MM) approaches have provided a general scheme for studies of chemical processes in proteins.[1−12] Significant progress has been made with calibrated semiempirical QM/MM approaches[2,7,10,11] that include careful evaluations of the relevant activation free energies by free energy perturbation approaches that date back to the 1980s.[13] These studies exploit the rapid evaluation of the semiempirical energies, and sample the phase space of the QM atoms and the surrounding MM atoms. However, the current challenge is to move to an ab initio representation with a QM/MM treatment, because such QM(ai) representations have been shown to provide "chemical accuracy" in studies of gas-phase reactions of small molecules. Here, we use the term ab initio mainly to differentiate form semiempirical methods. Therefore, QM(ai) may also refer to density functional theory, which uses empirical parameters in common implementations. Unfortunately, it is at present extremely challenging to evaluate the potential of mean force (PMF) for enzymatic reactions by QM(ai)/MM approaches[14,15] due to the requirement of very extensive sampling, which results in the extremely computationally expensive repeated evaluation of the QM energies.

The recent realization of the importance of the proper sampling of QM(ai)/MM surfaces led to several advances.[16−25] A major direction of these advances have been based on different adaptations[20−25] of the idea[16,17] of using a classical potential as a reference for the QM/MM calculations. Other strategies have also been quite promising.[19,26−28] Nevertheless, there is clearly a need for more "mainstream" approaches that can be used in standard implementations and aid in obtaining converging QM(ai)/MM free energies.

Here, we consider a simple and powerful treatment that can be viewed as a variable time step approach (from the QM/MM perspective) that appears to provide a powerful way to evaluate electrostatic free energies by QM/MM approaches. This ap-

* Corresponding author. E-mail: warshel@usc.edu. Phone: (213) 740 7671. Fax: (213) 740 2701.
† University of Southern California.
‡ National Institutes of Health (NIH).
§ University of Gdańsk.

10.1021/jp711496y CCC: $40.75
Published on Web 04/16/2008

proach is formally equivalent to approaches that add the average potential to the solute Hamiltonian and is thus a mean field approximation. Obviously, adding the average potential is an old idea that was implicitly implemented in the QM/Langevin Dipole (QM/LD) model.[29−31] It is also implemented implicitly in continuum models.[32−35] Furthermore, an averaging approach was implemented recently in an instructive work of Yang and co-workers.[26] However, while the addition of the average potential to the semiempirical Hamiltonian is very simple,[29] it requires specialized implementation in standard commercial QM(ai)/MM codes (that are designed to handle external point charges). Furthermore, the use of the average potential may not reproduce the average energy obtained by using the instantaneous potential in each time step, as in the common case when the solvent fluctuations are significant.

In some respects, our approach is close in spirit to the work of Aguilar and co-workers[36−39] who examined important aspects of this issue. However, our work focuses on examining the relationship between the QM/MM approaches for the evaluation of free energies that do not involve any averaging and of different averaging strategies. This issue presents a general problem that, to the best of our knowledge, has not yet been addressed. We also pay special attention to the contribution of the solute polarization that reflects the response to the potential from the surrounding solvent. Obviously, it is quite simple to estimate this effect by classical approximations,[30] but the uncertainties in such approximations can be quite significant in the challenging case of charged transition states in proteins.

The present approach maps the effect of the fluctuating MM environment during m time steps, on a set of $m \times L$ point charges (where L is the number of atoms within a relatively small cutoff radius) and two additional point charges that represent the rest of the environment. This seemingly simple implementation is found to be very effective, leading to computational time saving of a factor of 1000 in QM(ai)/MM calculations of solvation free energies where the solute structure is fixed. Furthermore, the present study casts an interesting light on the relationship between the solvent fluctuations and the convergence of QM(ai)/MM calculation. This insight will be exploited in QM(ai)/MM free energy calculations that allow the solute to fluctuate.

II. Methods

Our starting point is the QM/MM energy of our system that is expressed here as

$$E_{\mathrm{tot}} = \langle \phi_{\mathrm{s}} \psi_{\mathrm{S}}^{\mathrm{pol}} | H^{\mathrm{g}} + V_{\mathrm{Ss}}^{\mathrm{el}} + V_{\mathrm{Ss}}^{\mathrm{vdW}} + H_{\mathrm{ss}} | \phi_{\mathrm{s}} \psi_{\mathrm{S}}^{\mathrm{pol}} \rangle$$
$$\cong E_{\mathrm{QM}}^{\mathrm{pol}}(\mathbf{R},\mathbf{Q}(U_{\mathrm{s}})) + E_{\mathrm{QM/MM}}^{\mathrm{el}}(\mathbf{R},\mathbf{r},\mathbf{Q}(U_{\mathrm{s}})) + E_{\mathrm{vdW}}(\mathbf{R},\mathbf{r}) + E_{\mathrm{MM}}(\mathbf{r}) \quad (1)$$

where $\mathbf{R}$ and $\mathbf{r}$ are the solute (S) and solvent (s) coordinates, respectively, Ψ_{S} and ϕ_{s} are the wavefunction of the solute and solvent, respectively, and $\mathbf{Q}$ is the vector of solute residual atomic charges that depends on the potential created by the solvent (U_{s}). Here the first term, $E_{\mathrm{QM}}^{\mathrm{pol}}(\mathbf{R},\mathbf{Q}(U_{\mathrm{s}})) = \langle \Psi_{\mathrm{S}}^{\mathrm{pol}} | H_{\mathrm{g}} | \Psi_{\mathrm{S}}^{\mathrm{pol}} \rangle$, is the energy of the polarized gas-phase Hamiltonian with a solute wavefunction that reflect polarization by the solvent. The second term ($E_{\mathrm{QM/MM}}^{\mathrm{el}}(\mathbf{R},\mathbf{r},\mathbf{Q}(U_{\mathrm{s}}))$) is the solute−solvent electrostatic interaction. The third term, $E_{\mathrm{vdW}}(\mathbf{R},\mathbf{r})$, is the solute−solvent van der Waals interaction and the last term, $E_{\mathrm{MM}}(\mathbf{r})$, is the solvent potential surface. One can study different types of problems with above E_{tot}, and here we will take the evaluation of the solvation free energy as an example. Our

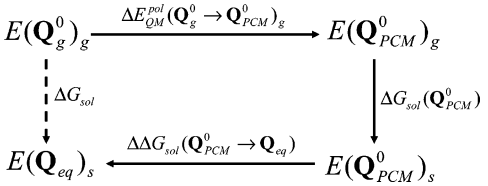

Figure 1. Energy scheme. $\Delta G_{\mathrm{sol}}(Q_{\mathrm{PCM}}^0)$ is calculated by the classical adiabatic charging approach.

starting point is the free energy perturbation (FEP) adiabatic charging (AC) approach[29,40] where we can use a mapping potential in the form

$$E_k = (E_{\mathrm{tot}} - E_{\mathrm{MM}})(1 - \lambda_k) + E'\lambda_k + E_{\mathrm{MM}} \quad (2)$$

where E' denotes energy of system without electrostatic solute−solvent interaction ($E' = \langle \Psi_{\mathrm{S}}^{\mathrm{g}} | H_{\mathrm{S}}^{\mathrm{g}} | \Psi_{\mathrm{S}}^{\mathrm{g}} \rangle + E_{\mathrm{vdW}}$) and where λ_k changes from zero to one in n + 1 steps.

We can use the standard FEP equation

$$\Delta\Delta G_{\mathrm{sol}}(\lambda_k \rightarrow \lambda_{k+1}) = -\beta^{-1} \ln\langle \exp\{-(E_{k+1} - E_k)\beta\}\rangle_{E_k}$$
$$\Delta G_{\mathrm{sol}} = \sum_{k=1}^{n+1} \Delta\Delta G_{\mathrm{sol}}(\lambda_k \rightarrow \lambda_{k+1}) \quad (3)$$

where $\beta = 1/(k_{\mathrm{B}}T)$; k_{B} is the Boltzmann constant, and T is the absolute temperature. The $\langle\rangle_{E_k}$ stands for the average obtained during the propagation of the configurations using the potential of E_k.

Equation 3 can be effectively approximated by using the linear response approximation (LRA) treatment[41]

$$\Delta G_{\mathrm{sol}} \cong \langle E_{\mathrm{tot}} - (E' + E_{\mathrm{MM}})\rangle_{E_{tot}} + \langle E_{\mathrm{tot}} - (E' + E_{\mathrm{MM}})\rangle_{E'} + \Delta G_{\mathrm{cav}} \quad (4)$$

where ΔG_{cav} is the solvation free energy of the nonpolar neutral form of the solute. This term consist of two parts describing hydrophobic and van der Waals free energy of cavity, which are not included in the first two terms of eq 4. They were described in ref 31 and implemented in ChemSol 2.0 program,[31] which was used to obtain ΔG_{cav} in this study. In the case of homogeneous solutions, the second LRA term of eq 4 is zero as the propagation of the configurations using the E' potential, which excludes the solute−solvent electrostatic interaction, does not lead to proper orientation of the solvent molecules.

Alternatively, we can combine the approaches of eqs 3 and 4 into a more efficient procedure presented in the following. This procedure uses the cycle of Figure 1 where we first run a classical MM simulation and use the FEP/AC approach[29] to evaluate the free energy of charging the solute to a given charge distribution (e.g., partial charges obtained with the PCM solvation model ($\mathbf{Q} = \mathbf{Q}_{\mathrm{PCM}}^0$, where "0" in superscript designates a constant value)), and then evaluate the change in the free energy allowing the charge to "equilibrate" with the solvent potential. The vector of equilibrated QM/MM residual atomic charges is designed by $\mathbf{Q}_{\mathrm{eq}}$. Thus the QM/MM solvation free energy can also be written as

$$\Delta G_{\mathrm{sol}} = \Delta E_{\mathrm{QM}}^{\mathrm{pol}}(\mathbf{Q}_{\mathrm{g}}^0 \rightarrow \mathbf{Q}_{\mathrm{PCM}}^0) + \Delta G_{\mathrm{sol}}(\mathbf{Q}_{\mathrm{PCM}}^0) + \Delta\Delta G_{\mathrm{sol}}(\mathbf{Q}_{\mathrm{PCM}}^0 \rightarrow \mathbf{Q}_{\mathrm{eq}}) \quad (5)$$

where the $\Delta G_{\mathrm{sol}}(\mathbf{Q}_{\mathrm{PCM}}^0)$ term is the solvation free energy of the solute, the atomic charges of which have been obtained from

PCM model. The $\Delta G_{sol}(\mathbf{Q}^0_{PCM})$ can be obtained using classical adiabatic charging approach based on the FEP treatment and therefore can be replaced by $\Delta G_{sol}(\mathbf{Q}=0 \rightarrow \mathbf{Q}^0_{PCM}) + \Delta G_{cav}$. The ΔE^{pol}_{QM} term in eq 5 is the polarization energy, which is given by

$$\Delta E^{pol}_{QM} = \langle \Psi^s_S | H^g_S | \Psi^s_S \rangle - \langle \Psi^g_S | H^g_S | \Psi^g_S \rangle \quad (6)$$

and $\Delta E^{pol}_{QM}(\mathbf{Q}^0_g \rightarrow \mathbf{Q}^0_{PCM})$ representing the polarization energy of a molecule in the PCM solvation model. The last term of eq 5, namely the $\Delta\Delta G_{sol}(\mathbf{Q}^0_{PCM} \rightarrow \mathbf{Q}_{eq})$, can be expressed using the LRA approach as

$$\Delta\Delta G_{sol}(\mathbf{Q}^0_{PCM} \rightarrow \mathbf{Q}_{eq}) \cong \frac{1}{2}[\langle E_{tot}(\mathbf{Q}) - E_{tot}(\mathbf{Q}^0_{PCM})\rangle_{E(\mathbf{Q})} + \langle E_{tot}(\mathbf{Q}) - E_{tot}(\mathbf{Q}^0_{PCM})\rangle_{E(\mathbf{Q}^0_{PCM})}] \quad (7)$$

where $E_{tot}(\mathbf{Q})$ is the QM/MM surface with the fluctuating charge. In case of both terms of eq 7, the calculation of the $E_{tot}(\mathbf{Q})$ and $E_{tot}(\mathbf{Q}^0_{PCM})$ is performed at exactly the same solute and solvent coordinates, and those systems differ only by the set of residual charges of the solute. Therefore, the average $< >$ term becomes the difference in the polarization energies and the interaction energies of both systems. The calculations of each $< >$ terms of eq 7 requires performing QM calculations as the configurations are propagated to obtain solute polarization energies included in $E_{tot}(\mathbf{Q})$. Therefore one can expect that calculation of eq 7 might be expensive.

Because $\langle \mathbf{Q} \rangle$ and $\mathbf{Q}^0_{PCM}$ are similar, the LRA treatment can be approximated by assuming that both terms are equal (this corresponds to a single step FEP with a very small $(E_0 - E_n)$ in eq 3). Thus, we can substitute the last term of eq 5 with an approximated eq 7, and write

$$\Delta G_{sol} = \Delta E^{pol}_{QM}(\mathbf{Q}^0_g \rightarrow \mathbf{Q}^0_{PCM}) + \Delta G_{sol}(\mathbf{Q}=0 \rightarrow \mathbf{Q}^0_{PCM}) + \Delta G_{cav} + \langle E_{tot}(\mathbf{Q}) - E_{tot}(\mathbf{Q}^0_{PCM})\rangle_{E(\mathbf{Q})} \quad (8)$$

Of course, we can also use eq 7 without approximating it any further.

The main time-consuming step in the evaluation of eq 8 is still the evaluation of the LRA term $\langle \rangle_{E(\mathbf{Q})}$ term, which is expensive even for studies of solvation free energies. Thus, we move to the principle point of the present paper, which is the acceleration of the evaluation of E_{tot}.

Obviously E_{tot} can be approximated by using fixed solute charge and writing

$$E_{tot} \cong E^g_{QM}(\mathbf{R}) + E^{el}_{QM/MM}(\mathbf{R},\mathbf{r},\bar{\mathbf{Q}}) + E_{vdW} + E_{MM} + \Delta E^{pol}_{QM}(\mathbf{Q}^0_g \rightarrow \bar{\mathbf{Q}}) \quad (9)$$

where $\bar{\mathbf{Q}}$ is some estimate of the solute charge in solution (from PCM or from QM results obtained at few solvent configurations). We can also replace $E^{el}_{QM/MM}$ by its classical approximation

$$(E^{el}_{QM/MM})_{cl} \equiv 332 \sum_{i(S)} \sum_{j(s)} \frac{Q^i q^j}{r_{ij}} \quad (10)$$

where q is the solvent residual charges, and i and j are indexes of the solute and solvent atoms, respectively.

ΔE^{pol}_{QM} can be estimated from LD or PCM calculations or even from a few QM/MM calculations at some solvent configurations. It is also possible to use the classical approximations considered in ref 30. Unfortunately, the approximations used for ΔE^{pol}_{QM} are very problematic because the solute wavefunction fluctuates during the solvent fluctuations. Here, the obvious alternative is to evaluate the polarization by using the averaged potential from the solvent. This can be done by adding the solvent averaged potential to the solute Hamiltonian that in the simplest semiempirical formulation will look like[29,30]

$$F_{\mu\mu} = F^{gas}_{\mu\mu} - \bar{U}_\mu \quad (11)$$

where F is the Fock matrix and μ is the given orbital on the μth atom and U_μ is the potential from the surrounding solvent on the μth atom. However, in the ab initio code the corresponding treatment is more complicated. Thus, although an innovative approach that adds the average solvent potential to the solute Hamiltonian was recently presented,[26] the most widely used codes treat the solvent potential by considering the solvent residual charge as an external charge, and our aim is to exploit this general feature in combining MM and QM programs.

Our approach, which is technically similar to the approach of Aguilar and co-workers,[36] is demonstrated schematically in Figure 1. In this strategy, we constrain the QM atoms (the solute atoms), evaluate the QM charges, $\mathbf{Q}^{(1)}$, where (1) designates the first step and run m MM/MD steps allowing the solvent molecule to move in the potential $(E^{el}_{QM/MM}(\mathbf{Q}^{(1)}) + E_{vdW})$. After m MD steps, we have m snapshots of solvent coordinates, which can be used in two ways. The first is to evaluate the average potential from the m snapshots but this would require a special implementation. The second option is to scale the charge of each solvent atom by $1/m$ and to send $m \times N$ solvent atoms with the scaled solvent charges to the QM program. This approach (Figure 2 and Figure 3b) is very simple but unfortunately it generates $m \times N$ external charges to be included into the Hamiltonian within the QM program. This can be too expensive and inconvenient, as will be shown in the following sections. Thus, we introduce the approximation described in Figure 3b. In this treatment, we divide the solvent into two regions. In the first region (region I), we convert the N_{ext} solvent atoms to $m \times N_{ext}$ external charges (scaled by $1/m$), while in region II, we represent the average solvent field coming from $N - N_{ext}$ solvent molecules, by two point charges (q and $-q$) using

$$\xi_O = \frac{2q}{|\mathbf{r}_{OR}|^3}\mathbf{r}_{OR} \quad (12)$$

where ξ_O is electric field at point O (the geometrical center of QM system) and $\mathbf{r}_{OR}$ is pointing along ξ_O to charge q.

The validity of the averaging approach is related to the approximation

$$\langle E_{tot}(\mathbf{r})\rangle^m = E_{tot}(\mathbf{Q}(\langle U\rangle^m)) \quad (13)$$

where $< >^m$ designates an averaged value over m MD steps. As will be seen below, it is not essential to satisfy eq 13.

The validity of the use of effective charges simply depends on the magnitude of region I. In this study, we consider four solvent representation models that reflect different sizes of region I and region II. Model M1 and model M4 are single layer models consisting only of region I or region II respectively. Models M2 and M3 are dual layers models and the corresponding sizes of region I and II are presented in Table 1. At any rate, our task is to explore the range of validity of the above approximations, and this will be done in the next section.

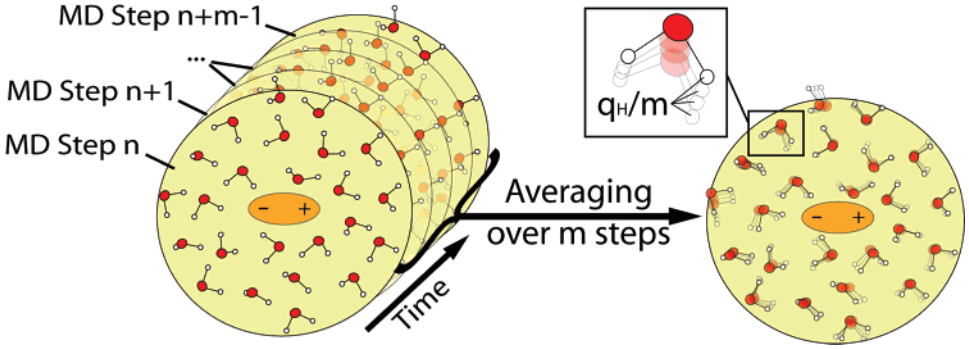

Figure 2. A schematic representation of the averaging of the solvent potential over m steps of a MD simulation.

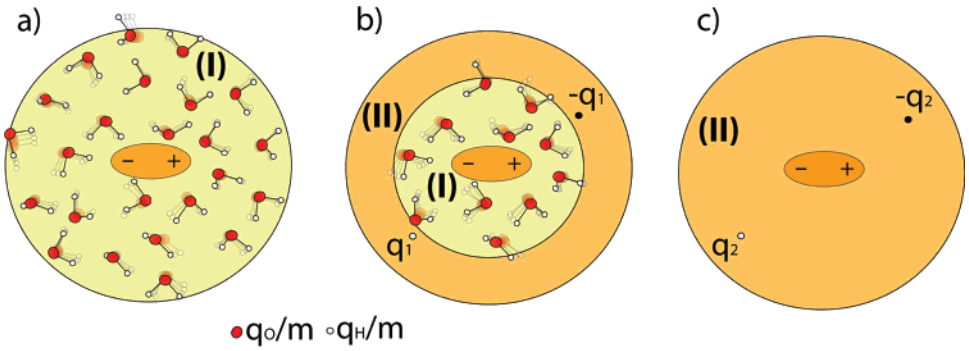

Figure 3. Different models for the evaluation of the average solvent charges: model M1 (a) involves an averaging of the explicit solvent molecules; models M2 and M3 (b) average the explicit molecules in region (I) while representing the average potential of the molecule in region II by two charges; model M4 (c) represents the solvent effect by two charges.

TABLE 1: The Different Solvent Representations Used in This Work

model name	radius of regions (I) and (II) in Å		number of explicit water molecules solvating different systems		comment
	r_I	r_{II}	H_2O	$HCOO^-$	
model M1	16	n/a	565	563	one layer, Figure 3a
model M2	10	16	137	143	two layers, Figure 3b
model M3	6	16	28	*a*	two layers, Figure 3b
model M4	n/a	16	0	0	one layer, Figure 3c

[a] This model was not used for this system.

The validation of the above approximations will be conducted for two different types of solutes: a polar molecule, water, and an $HCOO^-$ ion. All the solute atoms will be solvated by explicit water molecules placed in a sphere with a radius of 16 Å, which corresponds to 565 and 563 water molecules. respectively. All solvent molecules are represented by the ENZYMIX force field.[42] In the simulation model, the sphere of the explicit water molecules is surrounded by a surface region whose average polarization and radial distribution are determined by the surface-constrained all-atom solvent (SCAAS) model.[40,43,44] The surface region is embedded in a bulk continuum region with a dielectric constant of 80. The long range interactions are treated by the local reaction field (LRF) approach.[45]

The MD simulations presented here were performed using the MOLARIS package.[42] In every case, we first relaxed the system in a 10 ps long simulation of 1 fs time steps. All QM calculations were performed using the Gaussian 03 package. The B3LYP exchange-correlation potential[46,47] was used with 6-31+G** and aug-cc-pVDZ basis sets.[48] The Merz−Kollman scheme[49] with default atom radii was used to determine charges on atoms to be later used in the MD simulations.

The combined QM/MM calculations were facilitated by developing new extensions of MOLARIS that store solvent trajectories and calculate charges to be later used to define the external average potential in the QM calculation. The Gaussian 03−Molaris communication as well as free energy mapping was facilitated by application of the Gaussian Output Tools package[50] and other Perl scripts. All benchmarked calculations presented here were performed on dual Intel Xeon 3.06 GHz machines with 2 Gb of memory running the Red Hat Enterprise Linux 5 operating system.

III. Results and Discussion

III.1. A Solvated Water Molecule. In the first stage of our study, we evaluated the performance of the four different models (M1−M4) for representing the solvent effects on the solute while using the standard QM/MM approach ($m = 1$). This step was needed because the high m values ($m > 50$) are only accessible for models with a reduced number of explicit water molecules (models M2−M4). Figure 4 examines the averaged solute−solvent interaction energy ($<E_{int}>$) during a simulation of a solvated H_2O molecule in water using models M1−M4. In the case of the reference model (model M1), the value of $<E_{int}>$ converged after about 5 ps; here we define the convergence by the reaching of a constant value. The other models converged in a similar time (see below). The main point that emerged from the Figure 4 is that model M2 (where about 75% of explicit water molecules are represented by two charges) provided an excellent approximation: the plot of $<E_{int}>$ is practically parallel to that of model M1. The shift of ca. 0.5 kcal/mol originates in the beginning of the simulation (first 1.5 ps) and can be easily corrected by not taking the first 2ps of the QM/MM simulation (additional equilibration, see Table 2) into account. As seen from Figure 4 even the simplest models (M3 and M4) perform quite well, but they need much longer simulation times to converge. The longer simulation times can

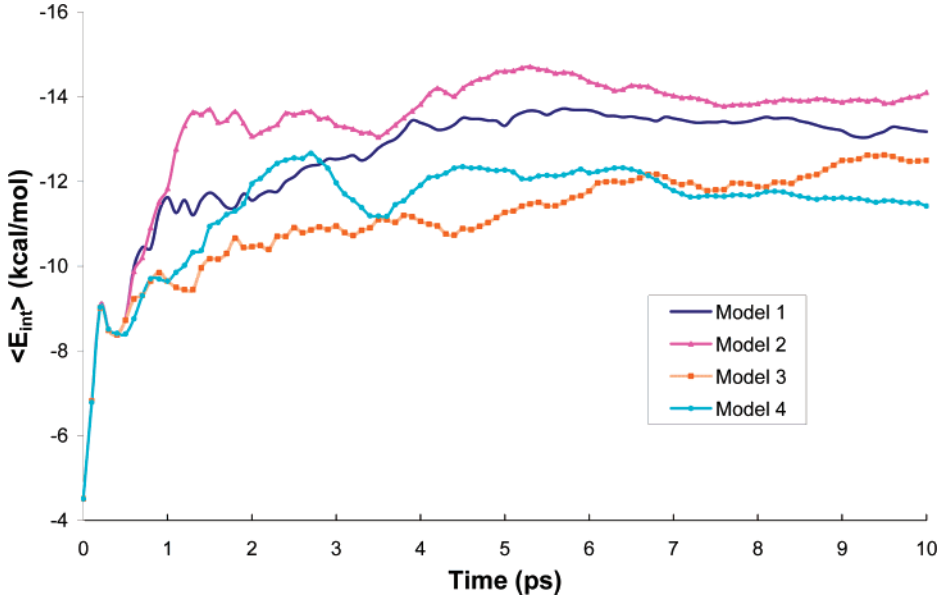

Figure 4. Convergence of the average interaction energy during 10 ps simulations ($m = 1$) using different models of solvent representation.

TABLE 2: Average Polarization Energy ($<E_{pol}>$), Average Solute−Solvent Electrostatic Interaction Energy ($<E_{int}>$), and Free Energy of Solvation (ΔG_{sol} of Equation 7) Obtained during the Last 8 ps of 10 ps Long Simulations of Water Molecule in a Water Solution[a]

	solvent representation models														
	model M1			model M2				model M3				model M4			
m	$<E_{pol}>$	$<E_{int}>$	ΔG_{sol}	$<E_{pol}>$	$<E_{int}>$	ΔG_{sol}	ΔG_{sol}*	$<E_{pol}>$	$<E_{int}>$	ΔG_{sol}	ΔG_{sol}*	$<E_{pol}>$	$<E_{int}>$	ΔG_{sol}	ΔG_{sol}*
1	1.6	−14.0	−6.1	1.6	−14.0	−6.1	*b*	1.4	−13.2	−5.9	*b*	1.0	−11.7	−5.3	*b*
10	1.4	−12.8	−5.8	1.4	−13.3	−5.9	*b*	1.4	−13.1	−5.8	*b*	1.1	−12.4	−5.7	*b*
25	1.4	−13.0	−5.7	1.9	−17.0	−7.5	−6.1	1.1	−11.2	−4.9	−5.9	1.4	−14.3	−6.3	−6.1
50	1.5	−14.2	−6.2	1.4	−13.5	−6.1	−6.3	1.5	−14.4	−6.4	−6.2	1.4	−15.4	−6.8	−6.3
100	*c*	*c*	*c*	1.5	−14.5	−6.4	−5.8	1.3	−13.6	−6.2	−6.4	1.1	−12.8	−6.0	−6.1
200	*c*	*c*	*c*	1.1	−11.7	−5.3	−6.0	1.3	−13.9	−6.2	−6.2	1.2	−14.0	−6.5	−6.3
500	*c*	*c*	*c*	1.1	−12.6	−5.7	−6.3	1.1	−12.5	−5.5	−6.0	0.8	−10.9	−4.9	−6.2
1000	*c*	*c*	*c*	1.3	−14.0	−6.5	−6.4	1.0	−11.7	−5.3	−5.9	0.8	−12.3	−5.8	−6.4

[a] The best estimates of the free energy of solvation (ΔG_{sol}* also of eq 7) are obtained by performing 50 and 100 ps simulations for $m = (25, 50)$ or $m = (100-1000)$, respectively. The free energies of solvation include the ΔG_{cav} term of 0.9 kcal/mol. [b]Not included in this study. [c]Cannot be completed due to insufficient memory of the available computer.

TABLE 3: Total QM/MM Simulation Time (Wall Time, τ, in Seconds) and Average Wall Time Required to Complete 1 QM Step (t) Obtained for 10000 Steps Simulation of Solvated a Water Molecule Using Solvent Potential Averaged over m Steps (Difference Solvent Representations Models Are Defined in Table 1)

	solvent representation model							
	model M1		model M2		model M3		model M4	
m	τ	*t*	τ	*t*	τ	*t*	τ	*t*
1	196941	22	192692	19	190123	19	189782	19
10	29816	30	21665	20	20568	19	20347	19
25	25980	62	10764	24	10103	19	8911	19
50	35502	170	7797	32	7092	29	5095	19
100	*a*	*a*	7536	61	4385	31	3184	19
200	*a*	*a*	9925	170	3125	36	2256	19
500	*a*	*a*	19436	893	2695	65	1657	19
1000	*a*	*a*	34803	3296	3119	168	1465	19

[a] Cannot be completed due to insufficient memory of the available computer.

only be reached by increasing *m*. However, as will be shown below, the converged results are practically the same as ones obtained with the reference model M1. The rationale for using models M2−M4 is clarified in Table 3, where we evaluate the timing of one QM calculation using each model for large values of *m*. In this study, models M2 and M4 are tested most extensively. Apparently, model M2 provides a good balance between calculation time and the simplification of the solvent representation. Model M4 has the big advantage of short calculation times (although it might require more steps to converge), while maintaining good accuracy. At the same time, its simplicity may seem somewhat artificial for those who may prefer more complete models.

After validating our approach for obtaining $\bar{U}_m$, we moved to the examination of the performance of eq 12 for different values of *m*, (note that our final focus will be on solvation free energies rather than eq 13.). This was done for the test case of a solvated water molecule in water where we evaluate major contribution to solvation free energy (solute−solvent electrostatic interaction energy (E_{int}) and solute polarization energy (E_{pol})) and dipole moment as a function of the simulation time.

Figure 5 describes the interaction energy for the different values of *m* as a function of time, while Figures 6 and 7 represent, respectively, the average solute−solvent electrostatic interaction energy and the solute polarization energy as a function of time. The plots represent 10 ps simulations using model M2. It can be seen from Figure 5 that the amplitude of E_{int} during the simulation is similar for the values of *m* studied (for the clarity of the plot only selected *m* values were presented in Figure 5). This is because *m* only defines how often the charges on the solute atoms are updated according to the average

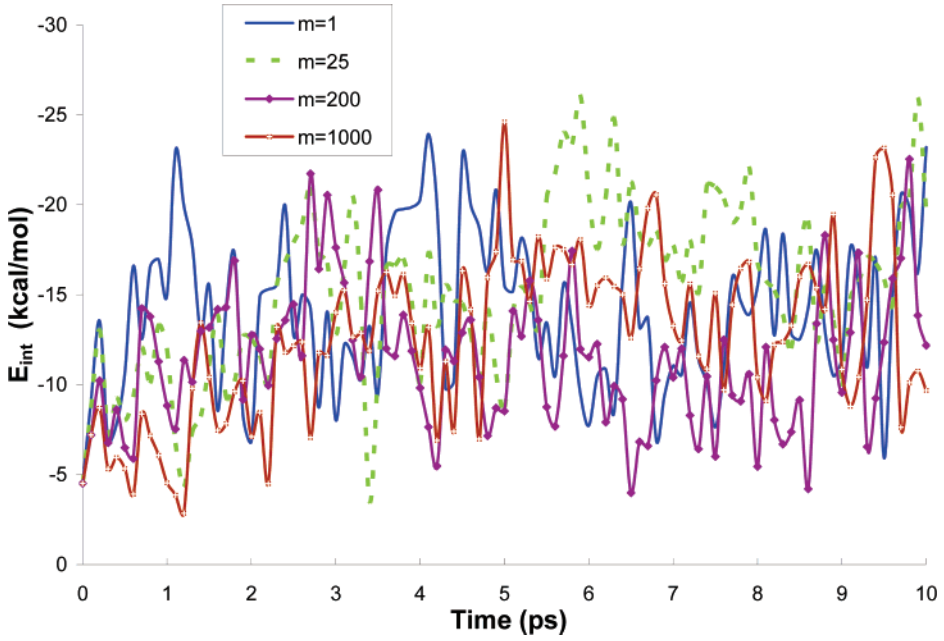

Figure 5. Interaction energy during 10 ps simulations of water molecule in water using model M2. The charge on the solute atoms were updated every *m*th step.

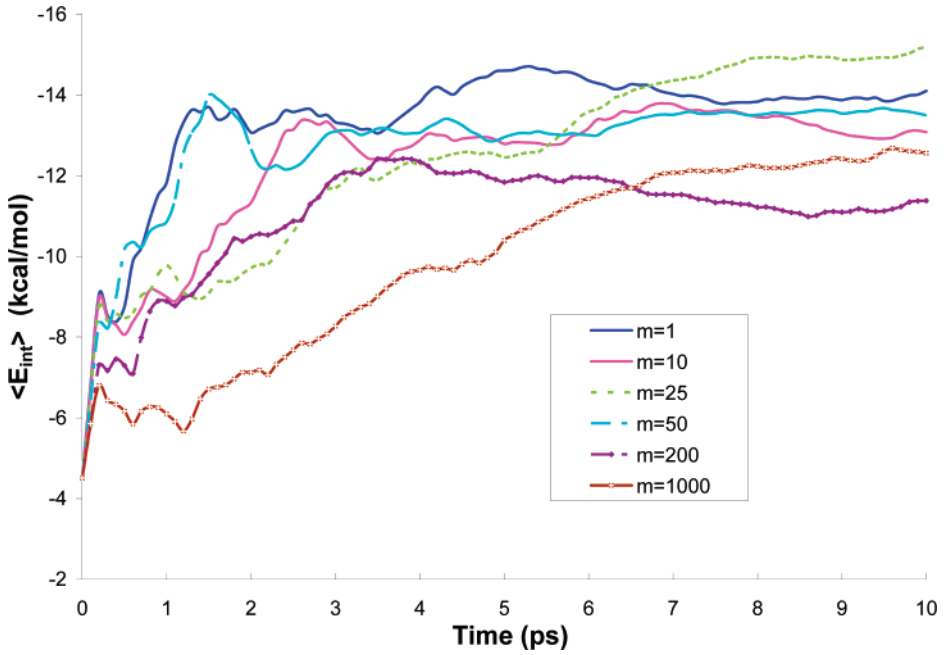

Figure 6. Average solute−solvent interaction energy during 10 ps simulations of a water molecule in water solution (model M2).

solvent potential. However, because the reasonable charges obtained in PCM calculations are used as the starting point of MD simulation, the changes in charges are not significant and it does not seem to affect the solute−solvent interaction to a great extent.

The plots of the averaged interaction energy, $<E_{int}>$, and the polarization energy, $<E_{pol}>$, lead to the conclusion that all selected m values lead to convergence. At the same time it is apparent that even the $m = 10$ simulation does not follow the $m = 1$ simulation in an exact way. These results might look surprising at first glance. However, a closer look should start from the realization that the solvent molecules fluctuate on a wide frequency range and that some of these frequencies involve periods which are shorter than the time between different QM/MM evaluations for $m > 1$. This point is clear from the Fourier transform of the fluctuating interaction energy (Figure 8), which includes oscillations above 800 cm^{-1} (about 7 fs). Fortunately, the fast oscillations do not contribute to the final free energy. Thus, we do not expect to have the same result for $m = 1$ and $m > 1$. In asking what the optimal value of m is, we have to realize that converging calculations require "equilibration" between the solute charge and the solvent potential. For example, we could use the approach described in Figure 9. In this approach, we run long MD simulation with $\mathbf{Q}^{(0)}$ and then take the resulting average potential, evaluate new $\mathbf{Q}$, and run another set of MD simulations. This procedure is repeated until convergence is reached. The plots demonstrating such convergence for different values of m and 3 models (M2−M4) are presented in Figures S-1−S-6 of the Supporting Information.

The convergence issue becomes very important when we evaluate solvation free energies by QM/MM models. Thus, we present the results for calculations of ΔG_{sol} for different values

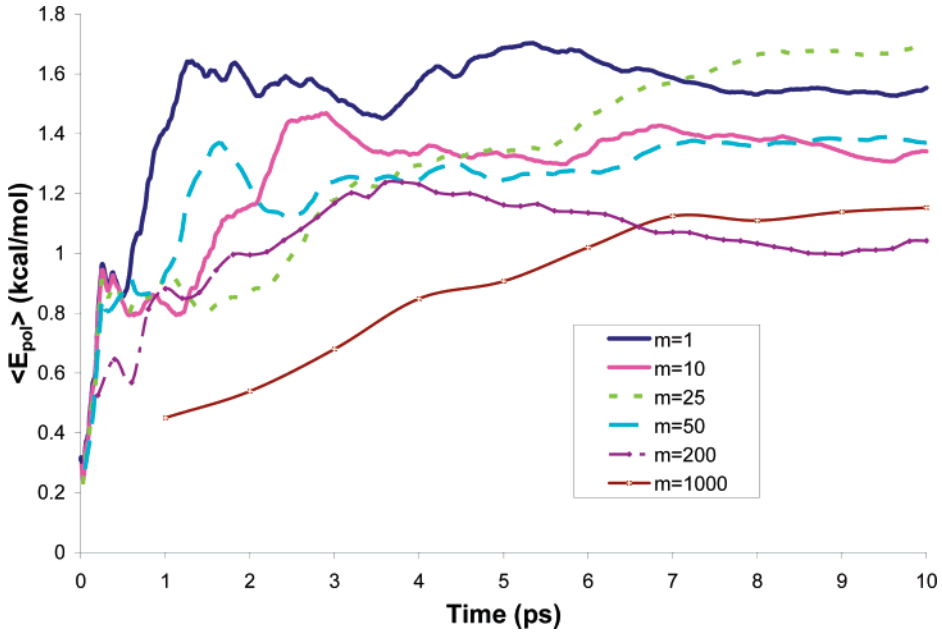

Figure 7. Average solute polarization energy during 10 ps simulations of water molecule in a water solution (model M2).

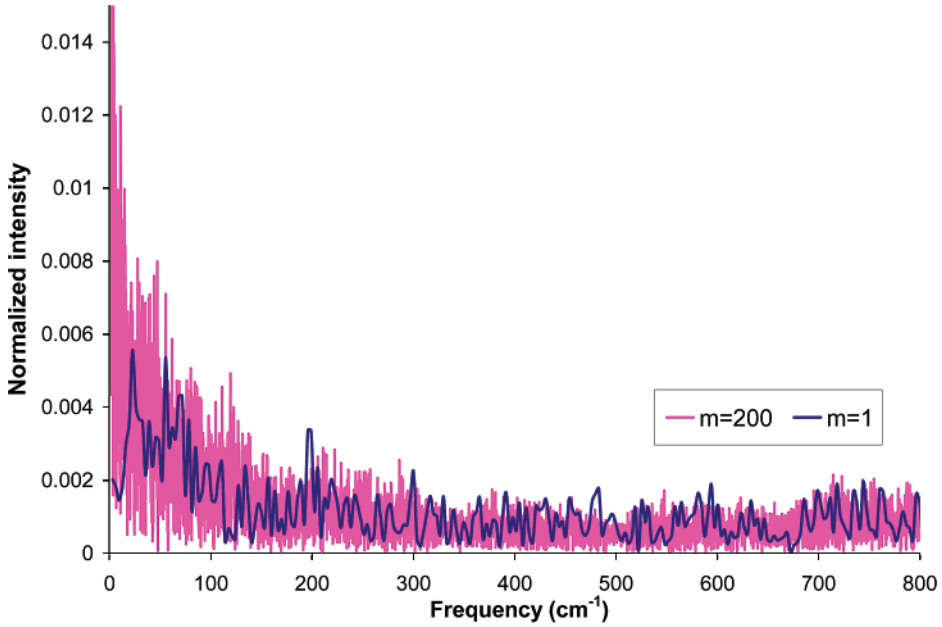

Figure 8. Fourier transform of the fluctuating interaction energy as a function of the simulation length and the value of *m* obtained in simulation of water molecule in water using model M2.

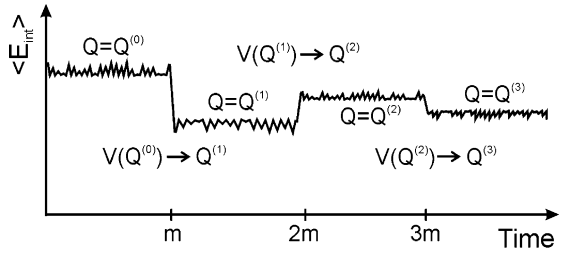

Figure 9. Schematic convergence.

of *m* and different solvent representations obtained after 10 ps simulation in Table 3. In the case of the model M1 and $m = 1$, the result of −6.1 kcal/mol is converged within a 10 ps simulation and is almost equal to experimental value of −6.3 kcal/mol.[51] When using the simpler solvent representation, model M2, the result is also converged and it the same as for model M1. For $m = 1$, models M3 and M4 give results further from the experimental value (−5.9 and −5.3, respectively). These results have not converged, but because of the computational cost, the longer simulation needed for full verification was not performed. Such longer simulations were performed for $m > 10$. As an example, let us consider the ΔG_{sol} results obtained with model M2 for $m = 25$ and $m = 200$. In the 10 ps simulation, we obtained −7.5 and −5.3 kcal/mol, respectively, which is up to 20% off the correct value. However, these results do not reflect proper convergence and longer simulation had to be conducted. The ΔG_{sol} results of 50 and 100 ps simulations conducted for $m = 25$ and $m = 200$, respectively, are in perfect agreement with the corresponding experimental result (Supporting Information Figures S1−S3 and Table 2). Similar results were obtained for model M1, which performed poorly over most of tested values of *m* in the 10 ps simulation. However, the free energy of solvation averaged over the 50 and 100 ps simulations reproduce the experimental value (Table 2).

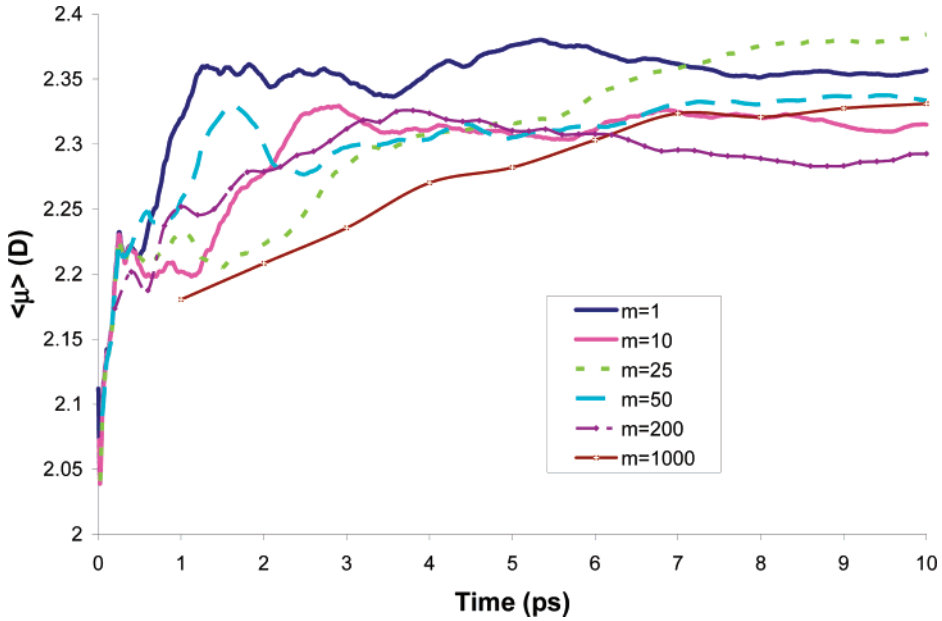

Figure 10. Average dipole moment of water molecule in water along a 10 ps simulation using model M2.

Interestingly, for all the solvent representation models presented here, $m = 50$ gives perfect agreement with the experimental result, but this is coincidental because $m = 50$ did not perform as well for other solvent representation models not presented in this study.

At this point, one should wonder what are our recommendations for choosing the optimal model for solvent representation and what is the optimal value of m for calculating the free-energy of solvation? As was shown here, there is strong coupling between these parameters and the required time to obtain converged results. When using models M1−M3, increasing m will lead to an increase of the single QM step calculation time (Table 3). This trend is drastic for model M1 and might also become evident for model M3 for very large values of m. The single QM calculation time using model M4 does not depend on m, although this model might require much more MD steps to converge. One also has to keep in mind that model M4 is only good for small molecules, as this model uses two charges to reproduce the electric field only in the geometrical center of the solute. Larger molecules might require more sophisticated approaches to reproduce the local polarization of the molecular fragments. Therefore, the selection of the value of m and the appropriate model should be based on the system being studied and the estimated total calculation time required for convergence. In the case of the water molecule ,the most efficient approach would be to perform 100 ps simulation using model M4 and $m = 1000$. In a comparison to the standard approach (model M1, $m = 1$), it offers calculation time savings by 100 times less, while still giving the same result. As mentioned before, chemists might prefer to use more complete solvent potential models like M2 or M3, which performed well for values of m from 25 to 1000 offering computational time savings by a factor of up to 73 in the presented case (M3, $m = 500$).

After exploring the energetic issue, it is useful to examine the convergence of the solute polarization, and this was done in the present study in terms of the dipole moment of water. As seen in Figure 10, the use of different values of m leads to the convergence of the dipole moment. The converged value of the dipole moment obtained with $m = 1$ is 2.35 D. For other values of m considered in this study, the values obtained are within 2.28−2.38 D, which correspond to shifts of 0.48−0.58 D compared to the gas-phase result. Other studies employing broad range of computational techniques suggested the shifts of 0.4−0.8 D.[52−54] Mendoza et al.[38] have used a mean field approach with averaging potential over a different number of configurations and obtained dipole moments between 2.72 and 2.78 D. At any rate, our aim is not to reproduce of the dipole moment, but rather compare the results obtained by different m values.

III.2. A Solvated Formate Ion. After studying a polar molecule (namely H_2O), we examined the performance of our approach in the case of solvated ion. As a test case, we consider the formate ion and the corresponding results summarized in Table 4 and Figures 12−15. The plot of the solute−solvent interaction energy, E_{int}, obtained in a number of simulations using $m = 1...1000$ and model M2 is presented in Figure 12. It can be seen from the figure that the amplitudes of E_{int} are on the same order of magnitude for all values of m considered. Also, the average solute−solvent interaction energy, $<E_{int}>$, appears to converge nicely (Figure 13). More specifically, $<E_{int}>$, converges (within 2 kcal/mol) in all simulations with $m < 100$. The convergence with $m = 1000$ appeared to be much slower because the solute charge was updated only 10 times and thus did not reach satisfactory equilibration. Additional plots demonstrating the convergence of $<E_{int}>$ in 50 and 100 ps simulations for $m = 25$, 50 and $m \geq 100$, respectively, are presented in Supporting Information Figures S-7 and S-8.

The most remarkable difference between the convergence of $<E_{int}>$ for H_2O and $HCOO^-$ is that the convergence is much slower in the latter case. This reflects the fact that the initial orientation of the solvent molecules is far from optimal in the $HCOO^-$ case and more MD steps are required to orient them according to electric field of an ion. As shown by the plots of $<E_{int}>$ along 50 and 100 ps simulations for $m = 25$, 50 and $m \geq 100$, respectively, the convergence within 1 kcal/mol is reached after 40 and 60 ps, respectively.

In addition to the solute−solvent interaction energy, we studied the polarization energy of the $HCOO^-$ ion. As shown in Figure 14, the average polarization energy, $<E_{pol}>$, converges within a 10 ps simulation, for all the values of m considered. The converged value spread between 3.5 and 5.5 kcal/mol. The lowest value of ca. 3.5 kcal/mol correspond to $m = 1000$ and is reflected in the low average solute−solvent interaction energy observed in this case.

TABLE 4: Average Polarization Energy ($<E_{pol}>$), Average Solute−solvent Electrostatic Interaction Energy ($<E_{int}>$), and Free Energy of Solvation (ΔG_{sol} of Equation 8) Obtained during the Last 8 ps of 10 ps Simulations of $HCOO^-$ Anion in Water Solution[a]

	solvent representation models									
	model M1				model M2					model M4
m	$<E_{pol}>$	$<E_{int}>$	ΔG_{sol}	t	$<E_{pol}>$	$<E_{int}>$	ΔG_{sol}	ΔG_{sol}*	t	ΔG_{sol}*
1	4.1	−156.2	−78.6	28	4.7	−160.9	−80.6	*b*	27	*b*
10	4.4	−158.7	−79.7	38	4.1	−157.8	−79.3	*b*	41	*b*
25	4.3	−159.1	−79.9	74	4.3	−159.3	−79.9	−80.1	45	*b*
50	4.0	−156.5	−78.9	175	4.2	−159.0	−79.8	−79.6	42	*b*
100	*c*	*c*	*c*	*c*	4.2	−157.0	−78.8	−79.8	73	−81.7
200	*c*	*c*	*c*	*c*	4.1	−157.6	−79.2	−80.9	174	−81.5
500	*c*	*c*	*c*	*c*	4.4	−163.7	−82.2	−80.3	918	−81.5
1000	*c*	*c*	*c*	*c*	3.8	−154.4	−77.8	−80.0	3379	−81.7

[a] t is the average time required for one QM step. The best estimates of the free energy of solvation (ΔG_{sol}*) are obtained by taking average over the last 80 ps of 100 ps simulations (available only for $m \geq 100$). Average energies and free energies are given in kcal/mol, a time is given in seconds. For all calculations using model M4, $t = 26$. The free energies of solvation include the ΔG_{cav} term of 1.5 kcal/mol. [b]Not included in this study. [c]Cannot be completed due to insufficient memory of the available computer.

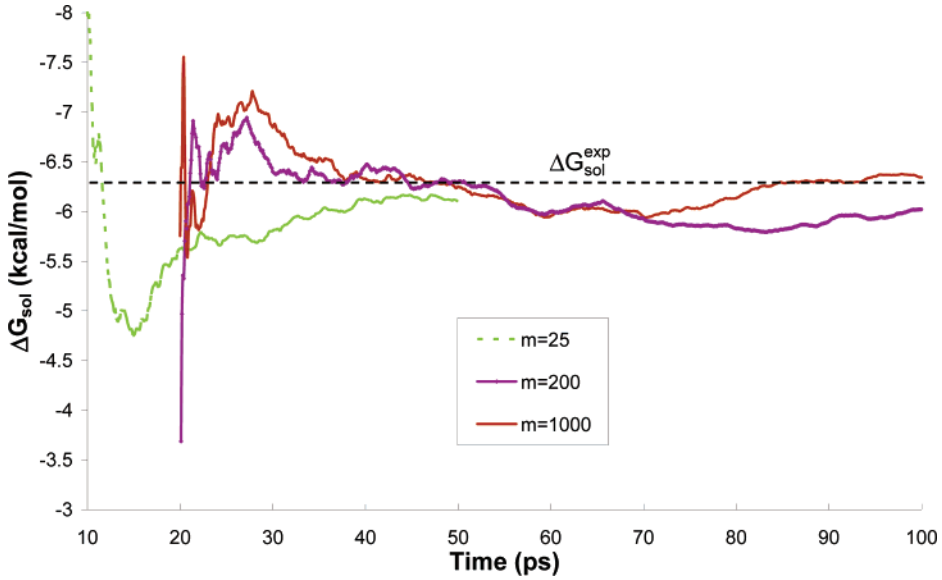

Figure 11. Free energy of solvation of water along 50 and 100 ps simulation of water molecule in water solution.

The converged result of $<E_{int}>$ and $<E_{pol}>$ can be used to obtain the solvation free energy using eq 4. The convergence of this approach can be assessed from Figure 15 (see the comment below). However, the final converged value turns out to be −88.0 kcal/mol, which represents a significant overestimation of the observed solvation energy (−80 kcal/mol[43]). A comparison of the classical ΔG_{sol} evaluated by the LRA approach of eq 4 and the FEP approach of eq 3 established that the overestimation is entirely due to the use of the LRA approximation for the entire charging process. Thus, we used eq 8 and obtained the results summarized in Table 4. As seen from the table, the calculated solvation free energy of $HCOO^-$ can converge to -80 ± 1 kcal/mol and is in excellent agreement with the corresponding observed value. Because the QM/MM part of eq 8 converges much faster than the corresponding part of eq 4, we describe in Figure 15 the convergence of eq 4 correcting the final result to those obtained by eq 8.

Overall the QM/MM treatment of eq 4 converges after about 40 and 60 ps simulations for $m = 25$, 50 and $m \geq 100$, respectively. On the other hand, the QM/MM part of eq 8 requires about 20 ps and the MM evaluation of $\Delta G_{sol}(\mathbf{Q}^0_{PCM})$ is very inexpensive relative to the QM/MM part.

To assess the effect of the model used, we also evaluate the solvation energies using model M4 (Table 4). The corresponding results are on average 1.5 kcal/mol lower than results obtained by the model M2 and even further from the experimental results.

Although the calculated free energy of solvation is in very good agreement with the experiment value, one has to keep in mind that the result can be inaccurate due to the QM method employed. The major contribution might be related to the size of the basis set used in QM calculations. We estimated these by conducting one test calculation (100 ps simulation with $m = 1000$ and model M4) where we employed the larger aug-cc-pVDZ basis set rather than 6-31++G** used before. The larger basis set is expected to better reproduce the polarization effect as it provides more flexibility for the SCF procedure. Indeed the calculated free energy of solvation was decreased by only less than 1% (0.7 kcal/mol).

When discussing the performance issues one should consider the purpose of the calculation. For example, when trying to obtain solvation energies by eq 4 one should consider using the M1 model, averaging over $m = 25-50$ solvent configurations and using accurate QM methods. In our test case of the $HCOO^-$ ion, the single QM calculation using model M1 and 50 solvent configurations was on average 6.35 times slower than calculation for $m = 1$ using the same M1 model. However, because 50 times less QM calculations are required the overall speed-up is 7.8. For studies that require obtaining relative results for the number of systems, one could use models M2−M3 and

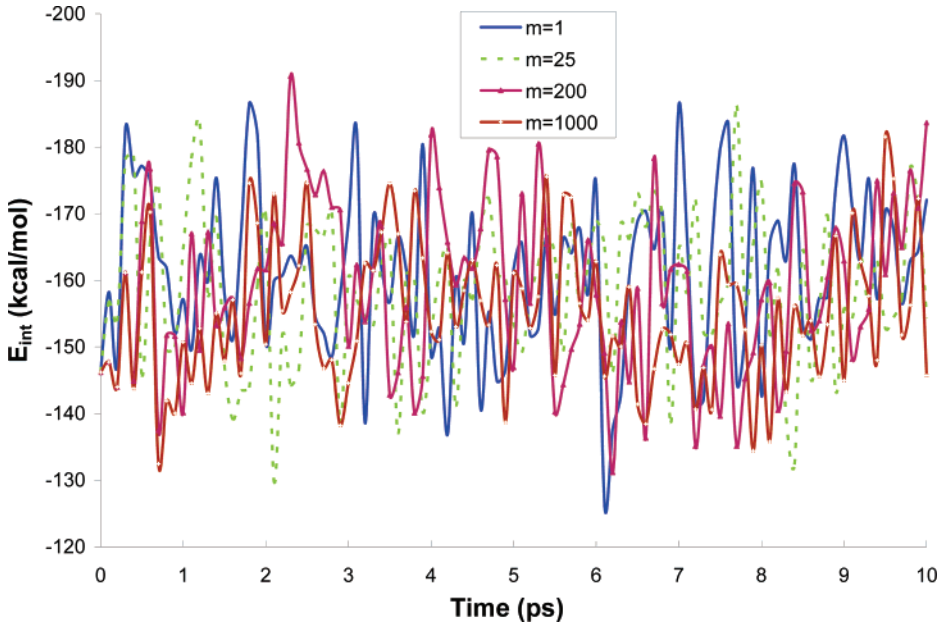

Figure 12. Interaction energy along 10 ps simulations of $HCOO^-$ in water using model M2. Charge on the solute atoms updated every *m*th step.

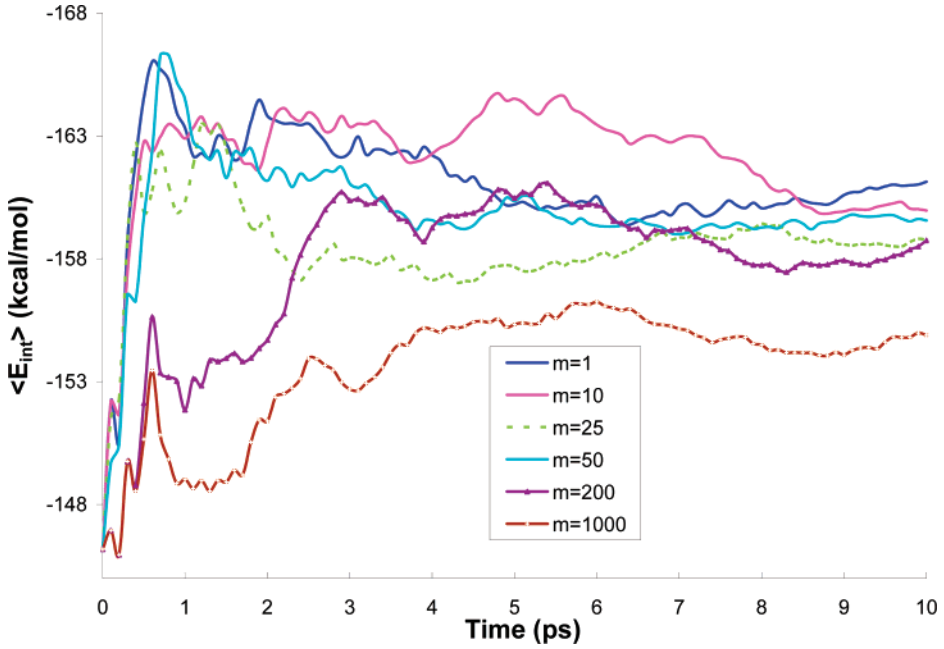

Figure 13. Average solute−solvent interaction energy along 10 ps simulation of $HCOO^-$ solvated by water (model M2).

$m > 100$, which provided convergence within 100 ps simulation requiring 100−1000 QM steps. Our benchmark for model M2 and $m = 200$ suggests that the single QM step is 6.5 times slower than for $m = 1$; however, with averaging over $m = 200$ the total speed-up using our accelerated approach is 30.6 times. On the other hand, for rough estimates of relative energies using the M4 model seems very attractive. In this case, the averaging even over 1000 steps takes as much time as single gas-phase calculation. Therefore, the total speed-up using our accelerated approach can reach 1000 times in this case. This of course does not take into account the slower convergence observed in the case of $HCOO^-$ for $m = 1000$, which in practice decreased the total speed-up to ca. 100 times, as a 10 times longer simulation was required.

IV. Concluding Remarks

The evaluation of the free energy changes in solutions and proteins by classical MM approaches usually requires very extensive averaging over the configurational space of the protein. This work establishes that the same requirement will hold for QM/MM calculations. Thus, it is clear that simple minimization approaches of the type used in gas-phase QM calculations would not be effective in evaluations of the activation energies of chemical reactions in proteins (e.g., see ref 55). Unfortunately, evaluating free energies of QM(ai)/MM surfaces is extremely challenging due to the need for extensive evaluation of the QM(ai) energies. Several innovative strategies for accelerating QM(ai)/MM sampling have already emerged.[16,17,19,26−28] Nevertheless, there is clearly a need for more "mainstream" ap-

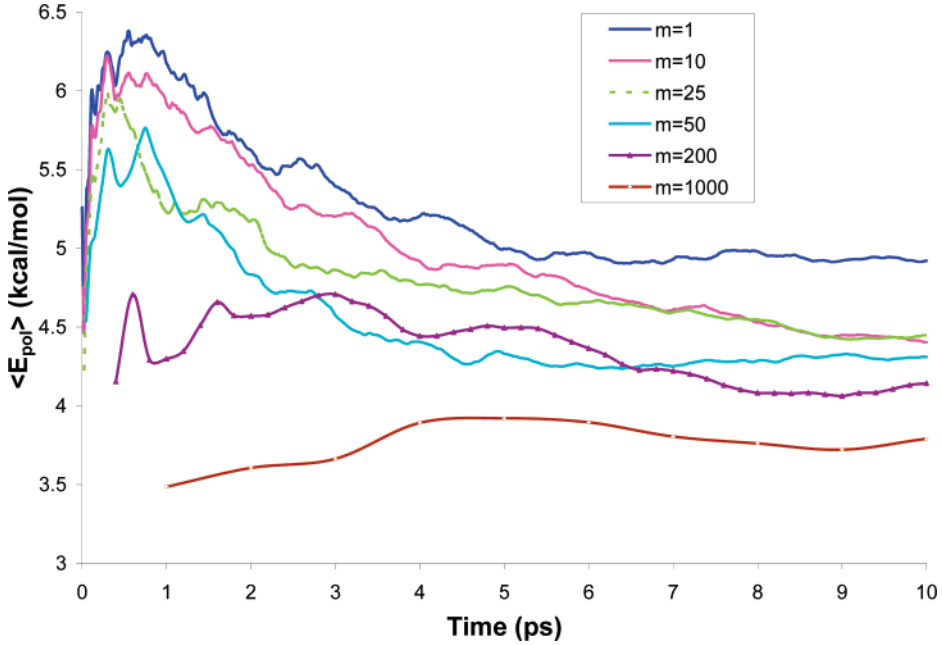

Figure 14. Average solute polarization energy along 10 ps simulation of $HCOO^-$ ion in water solution (model M2).

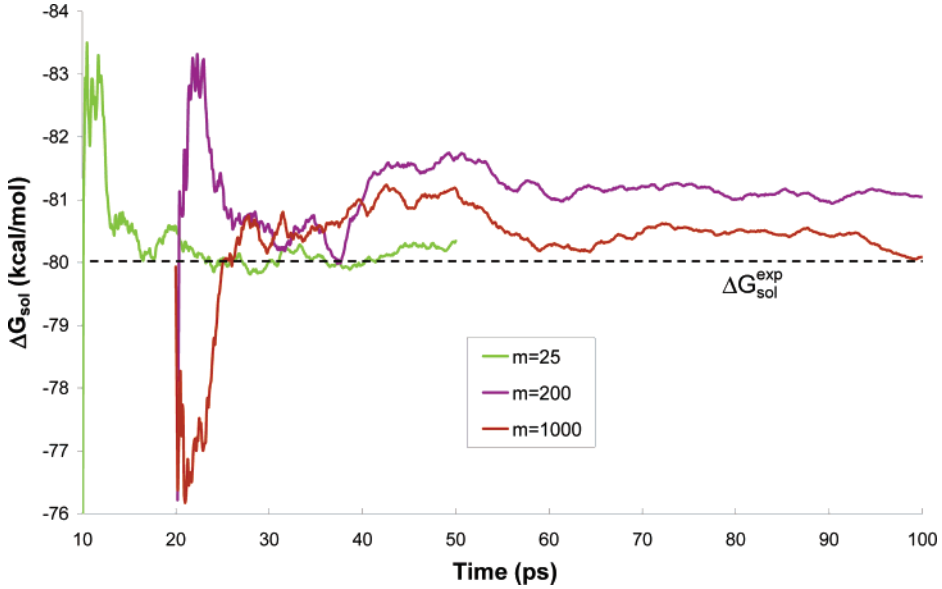

Figure 15. Free energy of solvation of the formate ion along 50 and 100 ps simulation. The calculations were performed using eq 4 and corrected by adding the difference between the ΔG_{sol} of eq 7 and ΔG_{sol} of eq 4.

proaches that can be used in standard implementations and help in obtaining converging QM(ai)/MM free energies.

The present study provides a practical and effective approach for performing QM(ai)/MM calculations of solvation free energies in solutions and in proteins. This is done by averaging the solvent potential over a fixed number of steps (m) in MM simulations, representing the effect of the fluctuating solvent charges by effective charges, and then updating the solute polarization by incorporating the effective charges in the solute Hamiltonian. The idea of using an average potential is obviously not new and has been used implicitly in our QM/LD models.[29−31] Furthermore, some important aspects of the current strategy have been used before. For example, the useful idea of representing the solvent potential by effective charges and a variable m value was used in refs 36−39, and an average potential was used in ref 26. However, the main point in our work is the systematic examination of the convergence of the averaging approach in actual free energy calculations and the relationship between the uses of different m values and the full QM/MM simulations that do not use any updating. Moreover, we explored the validity of using simple charge distributions that lead to significant savings in computer time. Finally, the key point of our approach is the development of an accurate and effective approach for calculating QM(ai)/MM solvation free energies.

Our study establishes the conditions for converging free energy calculations and provides several very effective ways of obtaining converged results, which to the best of our knowledge has not been accomplished to date. At any rate, our approach provides a time saving of up to 1000 in QM/MM calculations of solvation free energies while basically giving the same results that would be obtained in such calculations without averaging procedure. Of course, if one would run only 100 MD steps overall, one will not save much time. Establishing this point is quite important because the argument that one is using a mean field approximation usually implies that the results are approximated. However, we found here that the use of an

average potential in free energy calculations gives the same results as free energy calculations with the instant potential.

The present study has not considered the solvent polarization because it used the MOLARIS nonpolarizable force field. However, we would like to clarify that we have been using polarizable force fields in QM/MM calculations since 1976[1] and have been considering the solvent polarization in many works (for review on polarizable force-fields see ref 56). Thus, we can add the polarized solvent dipoles to the solvent average charges by using the polarizable MOLARIS force field. This detail is left, however, to subsequent works.

The approach used here for fixed solvent coordinates can be extended to the much more challenging (and arguably more important) case where both the solute and solvent are allowed to fluctuate. At present, there are only a few approaches that can be used to address this issue in an efficient way.[20−25] Thus, it is important to explore practical approaches for evaluating the PMF in the solute−solvent configurational space without fixing the solute coordinate during the free energy calculations. The most obvious use is the refinement of our approach that use the EVB as a reference potential.[15] In this approach, we perform MD simulation on both the EVB and corresponding QM/MM surface. Thus, we can obtain significant saving by reducing the expensive QM evaluations while using the average MM potential obtained between the number of QM evaluations. However, we also explore other less involved strategies. We believe that the present strategy will provide a useful way of advancing in this direction.

Acknowledgment. The authors would like to thank Prof. Peter Pulay for his comments on the implementation of distributed charges in the QM calculations and Dr. Shina Kamerlin for her comments of the manuscript. This work was supported by the NIH Grant GM24492. M.H. holds award from Foundation for the Development of the University of Gdansk (FRUG). The computational work was supported by University of Southern California High Performance Computing and Communication Center (HPCC).

Supporting Information Available: Figures S-1−S-8 are available as Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.

(2) Shurki, A.; Warshel, A. Structure/Function Correlations of Protreins using MM, QM/MM and Related Approaches; Methods, Concepts, Pitfalls and Current Progress. *Adv. Protein Chem.* **2003**, *66*, 249−312.

(3) Gao, J. Hybrid quantum and molecular mechanical simulations: an alternative avenue to solvent effects in organic chemistry. *Acc. Chem. Res.* **1996**, *29*, 298−305.

(4) Bakowies, D.; Thiel, W. Hybrid Models for Combined Quantum Mechanical and Molecular Approaches. *J. Phys. Chem* **1996**, *100*, 10580−10594.

(5) Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comp. Chem.* **1990**, *11*, 700−733.

(6) Friesner, R.; Beachy, M. D. Quantum Mechanical Calculations on Biological Systems. *Curr. Op. Struct. Biol.* **1998**, *8*, 257−262.

(7) Monard, G.; Merz, K. M. Combined Quantum Mechanical/Molecular Mechanical Methodologies Applied to Biomolecular Systems. *Acc. Chem. Res.* **1999**, *32*, 904−911.

(8) Garcia-Viloca, M.; Gonzalez-Lafont, A.; Lluch, J. M. A QM/MM study of the Racemization of Vinylglycolate Catalysis by Mandelate Racemase Enzyme. *J. Am. Chem. Soc* **2001**, *123*, 709−721.

(9) Marti, S. A., J.; Moliner, V.; Silla, E.; Tunon, I.; Bertran, J. *Theor. Chem. Acc.* **2001**, *3*, 207−212.

(10) Field, M. Stimulating Enzyme Reactions: Challenges and Perspectives. *J. Comp. Chem.* **2002**, *23*, 48−58.

(11) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. A QM/MM Implementation of the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) Method. *J. Phys. Chem. B* **2001**, *105*, 569−585.

(12) Lyne, P. D.; Mulholland, A. J.; Richards, W. G. Insights into Chorismate Mutase Catalysis From a Combined QM/MM Simulation of the Enzyme Reaction. *J. Am. Chem. Soc.* **1995**, *117*, 11345−11350.

(13) Warshel, A.; Sussman, F.; Hwang, J.-K. Evaluation of Catalytic Free Energies in Genetically Modified Proteins. *J. Mol. Biol.* **1988**, *201*, 139−159.

(14) Klahn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. On Possible Pitfalls in ab initio Quantum Mechanics/Molecular Mechanics Minimization Approaches for Studies of Enzymatic Reactions. *J. Phys. Chem. B* **2005**, *109*, (32), 15645−15650.

(15) Rosta, E.; Klahn, M.; Warshel, A. Towards Accurate Ab Initio QM/MM Calculations of Free-Energy Profiles of Enzymatic Reactions. *J. Phys. Chem. B* **2006**, *110*, (6), 2934−2941.

(16) Muller, R. P.; Warshel, A. Ab Initio Calculations of Free Energy Barriers for Chemical Reactions in Solution. *J. Phys. Chem.* **1995**, *99*, 17516−17524.

(17) Strajbl, M.; Hong, G.; Warshel, A. Ab-initio QM/MM Simulation with Proper Sampling: "First Principle" Calculations of the Free Energy of the Auto-dissociation of Water in Aqueous Solution. *J. Phys. Chem. B* **2002**, *106*, 13333−13343.

(18) Olsson, M. H. M.; Hong, G.; Warshel, A. Frozen Density Functional Free Energy Simulations of Redox Proteins: Computational Studies of the Reduction Potential of Plastocyanin and Rusticyanin. *J. Am. Chem. Soc.* **2003**, *125*, 5025−5039.

(19) Zhang, Y.; Liu, H.; Yang, W. Free energy calculation on enzyme reactions with an efficient iterative procedure to determine minimum energy paths on a combined *ab initio* QM/MM potential energy surface. *J. Chem. Phys.* **2000**, *112*, (8), 3483−3492.

(20) Rod, T. R. Quantum mechanical free energy barrier for an enzymatic reaction. *Phys. Rev. Lett.* **2005**, *94* (13), 138302.

(21) Liu, W. W. Doren, D. J. Density and temperature dependences of hydration free energy of Na+ and Cl- at supercritical conditions predicted by a initio/classical free energy perturbation. *J. Phys. Chem. B* **2003**, *107* (35), 9505−9513.

(22) Iftimie, R. S.; Schofield, J. An efficient Monte Carlo method for calculating ab initio transition state theory reaction rates in solution. *J. Chem. Phys.* **2003**, *119*, (21), 11285−11297.

(23) Crespo, A. M.; Estrin, D. A.; et al. Multiple-steering QM-MM calculation of the free energy profile in chorismate mutase. *J. Am Chem. Soc.* **2005**, *127*, 6940−6941.

(24) Sakane, S. Y.; Liu, W. B.; et al. Exploring the ab initio/classical free energy perturbation method: The hydration free energy of water. *J. Chem. Phys.* **2000**, *113*, (7), 2583−2593.

(25) Pradipta, B. Accelerating quantum mechanical/molecular mechanical sampling using pure molecular mechanical potential as an importance function: The case of effective fragment potential. *J. Chem. Phys.* **2005**, *122* (9), 091102.

(26) Hu, H.; Lu, Z. Y.; Yang, W. T. QM/MM minimum free-energy path: Methodology and application to triosephosphate isomerase. *J. Chem. Theor. Comput.* **2007**, *3* (2), 390−406.

(27) Ishida, T.; Kato, S. Theoretical Perspectives on the Reaction Mechanism of Serine Proteases: The Reaction Free Energy Profiles of the Acylation Process. *J. Am. Chem. Soc.* **2003**, *125* (39), 12035−12048.

(28) Ruiz-Pernia, J. J.; Silla, E.; Tunon, I.; Marti, S.; Moliner, V. Hybrid QM/MM potentials of mean force with interpolated corrections. *J. Phys. Chem. B* **2004**, *108* (24), 8427−8433.

(29) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons: New York, 1991.

(30) Luzhkov, V.; Warshel, A. Microscopic Models for Quantum Mechanical, Calculations of Chemical Processes in Solutions: LD/AMPAC and SCAAS/AMPAC Calculations of Solvation Energies. *J. Comp. Chem.* **1992**, *13*, 199−213.

(31) Florián, J.; Warshel, A. Langevin Dipoles Model for Ab Initio Calculations of Chemical Processes in Solution: Parametrization and Application to Hydration Free Energies of Neutral and Ionic Solutes, and Conformational Analysis in Aqueous Solution. *J. Phys. Chem. B* **1997**, *101*, 5583−5595.

(32) Tapia, O.; Goscinski, O. Self-consistent reaction field theory of solvent effects. *Mol. Phys.* **1975**, *29* (6), 1653−1661.

(33) Cramer, C. J.; Truhlar, D. G. Continuum Solvation Models: Classical and Quantum Mechanical Implementations. In *Reviews in Computational Chemistry*, Lipkowitz, K. B.; Boyd, D. B., Eds. VCH: New York 1995; Vol. 6, pp 1−72.

(34) Tomasi, J.; Bonaccorsi, R.; Cammi, R.; Delvalle, F. J. O. Theoretical Chemistry In Solution - Some Results And Perspectives Of The Continuum Methods And In Particular Of The Polarizable Continuum Model. *J. Mol. Struct.* **1991**, *80*, 401−424.

(35) Rivail, J. L.; Rinaldi, D. *Computational Chemistry: Review of Current Trends*: World Scientific Publishing: Singapore, 1995.

(36) Sanchez, M. L.; Martin, M. E.; Galvan, I. F.; del Valle, F. J. O.; Aguilar, M. A. Theoretical calculation of the Stark component of the solute-solvent interaction energy. Validity of the mean field approximation in the study of liquids and solutions. *J. Phys. Chem. B* **2002**, *106* (18), 4813−4817.

(37) Sanchez, M. L.; Martin, M. E.; Aguilar, M. A.; Del Valle, F. J. O. Solvent effects by means of averaged solvent electrostatic potentials: Coupled method. *J. Comput. Chem.* **2000**, *21* (9), 705−715.

(38) Mendoza, M. L. S.; Aguilar, M. A.; del Valle, F. J. O. A mean field approach that combines quantum mechanics and molecular dynamics simulation: the water molecule in liquid water. *J. Mol. Struct.* **1998**, *426*, 181−190.

(39) Sanchez, M. L.; Aguilar, M. A.; delValle, F. J. O. Study of solvent effects by means of averaged solvent electrostatic potentials obtained from molecular dynamics data. *J. Comput. Chem.* **1997**, *18* (3), 313−322.

(40) Warshel, A.; Sussman, F.; King, G. Free Energy of Charges in Solvated Proteins: Microscopic Calculations Using a Reversible Charging Process. *Biochemistry* **1986**, *25*, 8368−8372.

(41) Lee, F. S.; Chu, Z. T.; Bolger, M. B.; Warshel, A. Calculations Of Antibody Antigen Interactions - Microscopic And Semimicroscopic Evaluation Of The Free-Energies Of Binding Of Phosphorylcholine Analogs To Mcpc603. *Protein Eng.* **1992**, *5* (3), 215−228.

(42) Lee, F. S.; Chu, Z. T.; Warshel, A. Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by the POLARIS and ENZYMIX Programs. *J. Comp. Chem.* **1993**, *14*, 161−185.

(43) Warshel, A.; Russel, S. T. Calculations of electrostatic interactions in biological systems and in solutions. *Q. Rev. Biophys.* **1984**, *17*, 283−422.

(44) King, G.; Warshel, A. A Surface Constrained All-Atom Solvent Model for Effective Simulations of Polar Solutions. *J. Chem. Phys.* **1989**, *91* (6), 3647−3661.

(45) Lee, F. S.; Warshel, A. A Local Reaction Field Method for Fast Evaluation of Long-Range Electrostatic Interactions in Molecular Simulations. *J. Chem. Phys.* **1992**, *97*, 3100−3107.

(46) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(47) Becke, A. D. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372−1377.

(48) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron-Affinities Of The 1st-Row Atoms Revisited - Systematic Basis-Sets And Wave-Functions. *J. Chem. Phys.* **1992**, *96* (9), 6796−6806.

(49) Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived From Semiempirical Methods. *J. Comput. Chem.* **1990**, *11* (4), 431−439.

(50) Haranczyk, M.; Gutowski, M. Quantum mechanical energy-based screening of combinatorially generated library of tautomers. TauTGen: A tautomer generator program. *J. Chem. Inf. Model.* **2007**, *47* (2), 686−694.

(51) Cramer, C. J.; Truhlar, D. G. AM1-SM2 and PM3-SM3 parameterized SCF Solvation Models for Free Energies in Aqueous Solution. *J. Comput.-Aided. Mol. Des.* **1992**.

(52) Wei, D.; Salahub, D. R. A Combined Density Functional and Molecular Dynamics Simulation of a Quantum Water Molecule in Aqueous Solution. *Chem. Phys. Lett.* **1994**, *24*, 291.

(53) Tunon, I.; MartinsCosta, M. T. C.; Millot, C.; RuizLopez, M. F.; Rivail, J. L. A coupled density functional-molecular mechanics Monte Carlo simulation method: The water molecule in liquid water. *J. Comput. Chem.* **1996**, *17* (1), 19−29.

(54) Laasonen, K.; Sprik, M.; Parrinello, M.; Car, R. Ab-Initio Liquid Water. *J. Chem. Phys.* **1993**, *99* (11), 9080−9089.

(55) Klahn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. On Possible Pitfalls in ab initio QM/MM Minimization Approaches for Studies of Enzymatic Reactions. *J. Phys. Chem. B*, in press.

(56) Warshel, A.; Kato, M.; Pisliakov, A. V. Polarizable Force Fields: History, Test Cases and Prospects. *J. Chem. Theor. Comput.* **2007**, *3* (6), 2034−2045.