

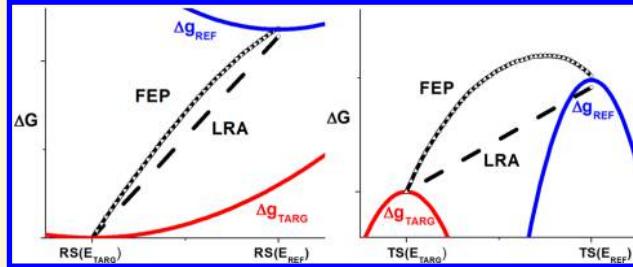
Exploring, Refining, and Validating the Paradynamics QM/MM Sampling

Nikolay V. Plotnikov and Arieh Warshel*

Department of Chemistry (SGM418), University of Southern California, 3620 McClintock Avenue, Los Angeles CA-90089, United States

 Supporting Information

ABSTRACT: The performance of the paradynamics (PD) reference potential approach in QM/MM calculations is examined. It is also clarified that, in contrast to some possible misunderstandings, this approach provides a rigorous strategy for QM/MM free energy calculations. In particular, the PD approach provides a gradual and controlled way of improving the evaluation of the free energy perturbation associated with moving from the EVB reference potential to the target QM/MM surface. This is achieved by moving from the linear response approximation to the full free energy perturbation approach in evaluating the free energy changes. We also present a systematic way of improving the reference potential by using Gaussian-based correction potentials along a reaction coordinate. In parallel, we review other recent adaptations of the reference potential approach, emphasizing and demonstrating the advantage of using the EVB potential as a reference potential, relative to semiempirical QM/MM molecular orbital potentials. We also compare the PD results to those obtained by direct calculations of the potentials of the mean force (PMF). Additionally, we propose a way of accelerating the PMF calculations by using Gaussian-based negative potentials along the reaction coordinate (which are also used in the PD refinement). Finally, we discuss performance of the PD and the metadynamics approaches in *ab initio* QM/MM calculations and emphasize the advantage of using the PD approach.



I. INTRODUCTION

Modeling chemical processes in condensed phases using the QM/MM method became the mainstream in computer simulations.^{1–5} Although recent methodological advances in electronic structure calculations and in parallel computing as well as an increase in available computer power, have opened new opportunities for predictive *ab initio* QM/MM simulations, evaluation of the free energies in *ab initio* QM/MM (QM(ai)/MM) studies has remained one of the major challenges in the field. One of the most promising options for performing QM(ai)/MM free energy calculations has been our approach^{6–8} of using a reference potential (RP) for the QM/MM potential and recent related versions discussed below. The most recent version of our idea of using the RP is implemented in the paradynamics (PD) model.⁹ The PD model is based in a large part on a number of earlier works and ideas including (a) those which have established the way of efficient sampling of the RP using the empirical valence bond¹⁰ (EVB) method combined with the evaluation of the free energy surface by a specialized combination^{11,12} of the free energy perturbation (FEP¹³) and umbrella sampling (US¹⁴) protocols, with the energy gap as a reaction coordinate;^{11,12} (b) those works which introduced the idea of using the EVB RP and the FEP/US approach in semiempirical⁷ and *ab initio*⁶ QM/MM free energy calculations; and (c) the ideas of refining the EVB RP⁸ to minimize the energy difference between the EVB RP

and the QM(ai)/MM target potential (TP), and using the linear response approximation^{15,16} (LRA) for calculating the FEP, while moving from the RP to the TP at the transition (TS) and reactant states (RS).

The PD approach has been found to be quite effective,⁹ where a typical evaluation of the activation free energy barrier by the PD approach (using the EVB RP) was estimated to require ~200 times less calls to the TP compared to the metadynamics (MTD) studies involving direct sampling of the TP.¹⁷ Thus the PD approach provides a rigorous way of dealing with the high computational cost of sampling of the QM(ai)/MM potential which is necessary to get an accurate free energy barrier.

In the present work, we report recent advances in expanding capabilities of the PD approach, clarifying possible misunderstandings and introducing several practical advances.

The first point that will be addressed is a misunderstanding associated with the fact that implementation of the PD scheme of moving from the RP to the TP has been mostly done with the LRA approach. This might have led some to the assumption that our approach is limited by validity of the LRA and thus does not yield an estimate of the relevant error. However, as we

Received: May 14, 2012

Revised: January 1, 2012

Published: August 1, 2012

tried hard to clarify, the perturbation between the two potentials can be done with a full FEP treatment. This point will be demonstrated here.

The second point that will be addressed in this work is our finding that the iterative PD approach⁹ of refining the EVB RP has not been fully robust. In particular, in some cases, one may want to change the EVB functional form rather than just refine the EVB parameters. In order to overcome this problem, we developed an approach that at least in some cases can be very effective in constructing and refining the RP in an alternative, more general way. This new approach involves fitting the potential energy scan (PES) along the specified reaction coordinate (RC) in the gas phase (or in an implicit solvent) for both the RP and the TP by Gaussian functions. The refined RP is constructed as the original RP modified with a correction that is taken as the difference between the Gaussian functions that approximate the TP and the RP.

The original RP approach⁷ has been adopted and implemented in a number of methods aimed at QM(ai)/MM calculations of activation free energies in enzymes^{18–20} and in aqueous solutions,²¹ as well as at evaluation of free energies of solvation,^{22,23} binding,²⁴ and conformational changes.²⁵ For example, in one recent approach, a semiempirical AM1/MM potential was used as the RP for the configurational sampling and for the construction of the reference free energy surface (FES). This was followed by an optimization on a DFT QM/MM potential at a number of points along the chosen RC, yielding the least energy reaction path. The entropic contribution was approximated by the difference between the activation free energy and the activation energy obtained on the AM1/MM RP. Furthermore, single point energy calculations on the optimized DFT QM/MM least energy path, using CCSD¹⁸ or MP2,¹⁹ were proposed to provide the activation free energy for a high level ab initio QM/MM surface. Another technique, introduced by Yang and co-workers,²⁶ and adopted elsewhere,²⁷ provides an estimate of the QM(ai)/MM free energy profiles with a fixed solute at a predetermined RC (which is determined, for instance, through iterative sequential optimization of the QM and MM regions independently, where in each step one of the regions is kept fixed while the second is optimized²⁶). This is followed by a single-step FEP from the MM-type RP to a QM/MM TP (with averaging on the RP). The reference FES is constructed by sampling on the RP. The ESP-derived solute charges for the MM-type RP are periodically updated by the QM/MM calculations. This method seems to be capable of capturing both the fixed solute polarization and the solvent polarization in a physically consistent way. A related approach involves single point energy evaluations^{20,28} with a high level QM/MM, performed at fixed solute geometries (in a similar way to that used in the work¹⁸ mentioned earlier). The energy difference between the DFT QM/MM and the high-level QM/MM was assumed to approximate the FEP while moving between these two potentials.²⁷ In another recent work, a different semiempirical RP, PM3/MM,²⁵ was used to construct the reference FES, followed by a single step FEP from the RP to a DFT QM/MM TP with the average being calculated on the RP.

The overall RP results, obtained in the above studies, clearly indicate that this approach is a powerful and versatile tool, which allows one to overcome the high cost of direct sampling on the QM(ai)/MM potential. However, there is a great potential for improvements to some of the existing RP treatments, especially considering the fact that some of them

overlooked our earlier, and arguably more consistent, RP treatment. For instance, one can foresee certain disadvantages to the use of minimization in exploring the multidimensional DFT QM/MM surface, as it is expensive and rather computationally inefficient. In addition to this, such a strategy can lead to multiple local minima. One has to keep in mind that the entropic contribution to the activation free energy barrier can be different for the TP and the RP, and that the activation energy is not equal to the activation enthalpy. Another serious problem is the slow convergence of the single step FEP approach when the average is taken only on the RP. This is especially problematic when the difference between the TP and the RP is significant (and their overlap in the phase space is small). In this case, sampling on the RP poorly represents the relevant region of interest on the TP.

The above problems are largely eliminated by the PD approach mainly due to the following features: (a) using the end points LRA that takes the averages both on the RP and on the TP and (b) using the PD refinement, which ensures faster LRA convergence due to a higher overlap between the TP and the refined RP. These points will be illustrated and discussed below.

An important element of the PD approach is the idea of refining the RP so it will be as close as possible to the TP. Our main refinement strategy starts with evaluation of the energy gap between the RP and the TP for the geometries generated by MD trajectories propagated at the RS and at the TS on the RP as well as on the TP. This step is followed by a least-squares minimization of the energy difference (plus other quantities). A somewhat similar refinement strategy has also been introduced in another context.²⁹ Although this approach is very effective, we found out that in cases when the improvement of the EVB requires a new functional we can get an improved refinement by adding Gaussian functions to the RP in a way that minimizes the difference between the RP and TP along the RC. A related idea of deforming the original potential to eliminate local minima was proposed by Scheraga and co-workers³⁰ in order to find the global minimum on a multidimensional surface. The use of a single negative Gaussian in the bias potential was also found to be useful in a MC study performed by Jorgensen and co-workers;³¹ however, it was for a rather stand-alone, specific case. A more general strategy of improved sampling by iteratively fitting the negative of the original potential with Gaussians in the local elevation method³² was formulated by van Gunsteren and co-workers. This idea was adopted, generalized in an elegant way, and popularized in the MTD approach of Parrinello and co-workers.³³ As will be clarified below, unlike these approaches, ours does not require expensive iterative sampling while building the negative potential. This step is accomplished by fitting the PES, and using the negative potential approximated by Gaussians in our FEP/US PMF³⁴ calculations. We accelerate convergence of PMF calculations with explicit solvent models by using the EVB type solvent driving potential, which forces the solvent to be polarized by the solvated solute charges along the reaction path.

In summary, we carry out two separate applications with the Gaussian functions: (a) during the PD refinement procedure, we evaluate two PESs (on the RP and on the TP), fit them both with Gaussian functions, and take the difference as a correction to the original RP; and (b) as the improved sampling strategy in FEP/US PMF calculations, we evaluate the negative of the original potential (without any iterative approaches) and use the corresponding Gaussians to create a flat mapping

potential. The sampling approach is augmented by the EVB-type solvent driving potential for the condensed phase calculations with explicit solvent models. At any rate, our Gaussian refinement strategy is as related to the MTD as our idea of building the RP that will be as similar as possible to the TP. As we clarified in our previous work,⁹ this early RP idea is formally identical to the MTD idea of building the negative of the TP but allows for more effective implementation.

II. METHODS AND RESULTS

II.1. The Standard Paradynamics Approach. The central idea of the PD (RP) approach is that the extensive configurational sampling required to calculate the QM(ai)/MM free energy barrier can be done on a computationally inexpensive reference potential^{6–8,16} rather than directly on the expensive target QM(ai)/MM potential. Our standard approach, described in Figure 1, involves construction of the

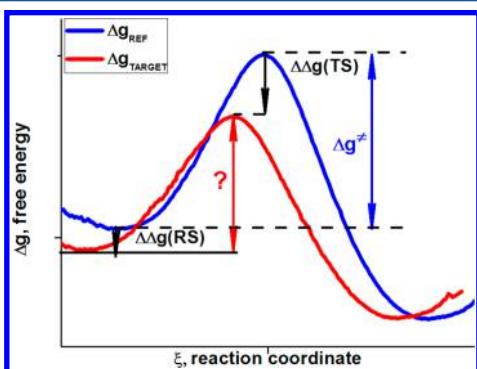


Figure 1. The thermodynamic cycle used in the paradynamics approach to calculate the free energy barriers. The blue curve is the free energy surface, $\Delta g(\xi)$, obtained for the reference potential with the corresponding barrier, Δg^\ddagger (blue arrow); the red curve is the free energy surface for the target potential, for which we want to evaluate the free energy barrier (red arrow). To achieve that, we estimate the free energy of moving from the reference potential to the target potential at the RS and at the TS (shown by black arrows).

free energy surface (FES) using the EVB¹⁰ method and the FEP/US technique.^{12,14} This is followed by evaluating the FEP using the LRA approach, while moving from the EVB RP to the TP at the RS and at the TS.

In considering activation barriers for chemical reactions, we follow our previous works¹⁶ and clarify the difference between the free energy profile (or the PMF), Δg , and the free energy of the system in a particular state, ΔG . This is done by starting with the expression for the rate constant, k :^{35,36}

$$k = \kappa \frac{\langle |\dot{\xi}| \rangle_{\text{TS}}}{2\Delta\xi^\ddagger} \frac{\Delta\xi^\ddagger \exp(-\beta\Delta g^\ddagger)}{\int_{-\infty}^{\xi^\ddagger} \exp(-\beta\Delta g(x)) dx} \\ = \kappa \frac{\langle |\dot{\xi}| \rangle_{\text{TS}}}{2\Delta\xi^\ddagger} \exp(-\beta\Delta G^\ddagger) \quad (1)$$

where κ is the transmission factor, $\dot{\xi}$ is the velocity along the RC, ξ , $\Delta\xi^\ddagger$ is the width of the TS region, Δg_{RS} is the value of the free energy profile (or PMF) at the minimum at the RS ($x = \xi_{\text{RS}}$), and Δg^\ddagger is its value at the TS ($x = \xi^\ddagger$). Here ΔG^\ddagger is defined by

$$\exp(-\beta\Delta G^\ddagger) = \frac{\exp(-\beta(\Delta g^\ddagger - \Delta g_{\text{RS}}))}{\frac{1}{\Delta\xi^\ddagger} \int_{-\infty}^{\xi^\ddagger} \exp(-\beta(\Delta g(x) - \Delta g_{\text{RS}})) dx} \quad (2a)$$

or

$$\Delta G^\ddagger = \Delta g^\ddagger - \Delta g_{\text{RS}} \\ + kT \ln \left(\frac{1}{\Delta\xi^\ddagger} \int_{-\infty}^{\xi^\ddagger} \exp(-\beta(\Delta g(x) - \Delta g_{\text{RS}})) dx \right) \quad (2b)$$

where the last term of eq 2b provides a convenient way of incorporating the entropic effect of the ground state into the TS-theory rate constant.³⁵ For the free energy of the RS, we also have

$$e^{-\beta\Delta G(\text{RS})} = \int_{-\infty}^{\xi^\ddagger} e^{-\beta\Delta g(x)} dx \quad (3)$$

This work will use the lower case $\Delta g(\xi)$ and the capital ΔG , as respectively, the value of the free energy profile, or the PMF, at a particular RC value (e.g., at TS, RS) and the free energy of a particular state (e.g., RS or TS defined by eq 3 or by $\Delta\xi^\ddagger$).

The second key idea of the PD approach, which drastically improves its computational cost, is the use of the LRA approach to calculate the free energy differences between the RP and the TP shown in Figure 1.

For example, in the case where the TS position is identical for the two free energy surfaces, we can write¹⁶

$$\Delta\Delta g_{E_{\text{REF}} \rightarrow E_{\text{TARG}}}(\xi_{E_{\text{TARG}}}^\ddagger) \\ = -kT \ln \frac{q_{E_{\text{TARG}}}(\xi_{E_{\text{TARG}}}^\ddagger)}{q_{E_{\text{REF}}}(\xi_{E_{\text{TARG}}}^\ddagger)} \\ = -kT \ln \left\langle \exp \left(-\frac{E_{\text{TARG}}(\xi_{E_{\text{TARG}}}^\ddagger) - E_{\text{REF}}(\xi_{E_{\text{TARG}}}^\ddagger)}{kT} \right) \right\rangle_{E_{\text{REF}}} \\ \cong \frac{1}{2} \left(\langle E_{\text{TARG}}(\xi_{E_{\text{TARG}}}^\ddagger) - E_{\text{REF}}(\xi_{E_{\text{TARG}}}^\ddagger) \rangle_{E_{\text{TARG}}} \right. \\ \left. + \langle E_{\text{TARG}}(\xi_{E_{\text{TARG}}}^\ddagger) - E_{\text{REF}}(\xi_{E_{\text{TARG}}}^\ddagger) \rangle_{E_{\text{REF}}} \right) \quad (4)$$

where the q 's are the partition functions at the specific RC values.

Finally, the third important component of the PD approach is the refinement of the RP, which brings it closer to the TP and ensures fast convergence of the FEP calculation from the RP to the TP. The above LRA expression also provides a framework for the PD refinement procedure, where the EVB parameters are refined by seeking the minimum of the least-squares function:⁹

$$\mathcal{T}(\mathbf{p}, \mathbf{r}) = k_1 \sum_{i=1}^N (E_i^{\text{EVB}}(\mathbf{p}, \mathbf{r}) - E_i^{\text{QM}}(\mathbf{r}))^2 \\ + k_2 \sum_{i=1}^N \sum_{j=1}^3 \left(\frac{\partial}{\partial x_j} E_i^{\text{EVB}}(\mathbf{p}, \mathbf{r}) - \frac{\partial}{\partial x_j} E_i^{\text{QM}}(\mathbf{r}) \right)^2 \quad (5)$$

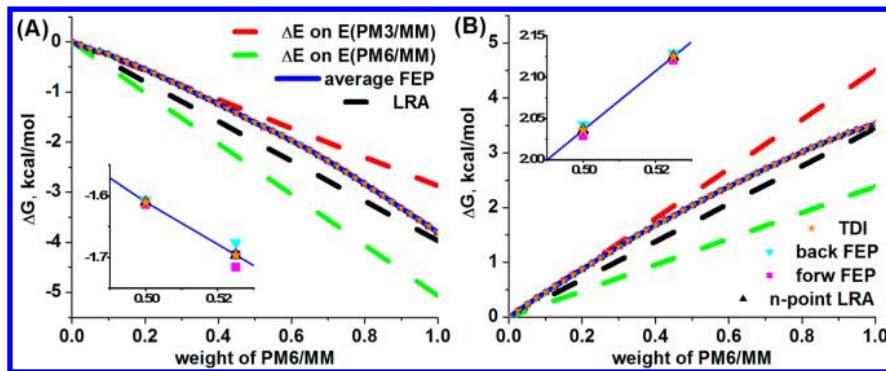


Figure 2. Comparative performance of the LRA approach and of a full multi-step FEP treatment in estimating the free energy of moving from $E_{\text{PM3/MM}}$ to $E_{\text{PM6/MM}}$ at the TS (A) and at the RS (B). The estimates were obtained by: 2-point LRA (black dashes); single-point LRA with the average of the energy gap, $E_{\text{PM6/MM}} - E_{\text{PM3/MM}}$, calculated on $E_{\text{PM3/MM}}$ (red dashes) and on $E_{\text{PM6/MM}}$ (green dashes); forward FEP (magenta squares); backward FEP (cyan triangles); average FEP (blue line); n -point LRA (black triangles); TDI (orange stars). Note that FEP-based estimates are hardly distinguishable due to overlap (they are shown in insets).

by either using a simplified Newton–Raphson approach where we refine one EVB parameter, p_i^i , at a time, where the result of the $k+1$ -th iteration is given

$$p_{k+1}^i = p_k^i - g[\partial\mathcal{T}/\partial p_i]/[\partial^2\mathcal{I}/\partial p_i^2]_k \quad (6)$$

or just by refining the vector of EVB parameters, \mathbf{p} , in the optimal steepest descent approach, where the result of the $k+1$ -th iteration is given by

$$\mathbf{p}_{k+1} = \mathbf{p}_k - s \cdot \text{grad } \mathcal{T}(\mathbf{p}_k) \quad (7)$$

The PD approach can be applied to any RP. However, the EVB potential has been the RP of choice, due to its very low computational cost, good performance in condensed phases (in particular in studies of enzymatic reactions¹⁶), physically consistent treatment of the solute–solvent electrostatic coupling, and other advantages discussed elsewhere.³⁷

While the refinement scheme given above is for EVB parameters, we also provide below a benchmark example of using a more expensive (compared to EVB) semiempirical RP to demonstrate that the PD approach is general as far as the RP is concerned.

II.2. Validating the LRA by a Full FEP Treatment.

II.2.1. The Case of Two General Potentials. Although the LRA treatment of eq 4 has been effective, its overall convergence may be problematic when the difference between the RP and the TP is significant. This has led some to assume that the validity of the LRA treatment limits application of our PD approach. However, what is missing in this assumption is the fact that LRA is just the end point approximation of the full FEP treatment which can be easily implemented by using

$$\Delta\Delta G_{\text{FEP}} = -kT \sum_{m=1}^{n-1} \ln \left\langle \exp \left(-\frac{E_{m+1} - E_m}{kT} \right) \right\rangle_{E_m} \quad (8)$$

where n is the number of simulation windows, between which the perturbation parameter λ (here weight of the TP) changes from 0 to 1, and E_m is the mapping potential given by

$$E_m = (1 - \lambda_m)E_{\text{REF}} + \lambda_m E_{\text{TARG}} \quad (9)$$

E_{REF} and E_{TARG} designate the RP and the TP, respectively. Better convergence is obtained, however, if we take the average of forward and backward FEP (see, e.g., ref 34):

$$\Delta\Delta G_{\text{FEP}} = \frac{1}{2} \left(\sum_{m=1}^{n-1} -kT \ln \left\langle \exp \left(-\frac{E_{m+1} - E_m}{kT} \right) \right\rangle_{E_m} - \sum_{m=2}^n -kT \ln \left\langle \exp \left(-\frac{E_{m-1} - E_m}{kT} \right) \right\rangle_{E_m} \right) \quad (10)$$

With the accurate results of the FEP method, we can explore the validity of the end point LRA, which gives (see also eq 4) the relevant $\Delta\Delta G$ by¹⁵

$$\Delta\Delta G_{\text{LRA}} = \frac{1}{2} (\langle E_{\text{TARG}} - E_{\text{REF}} \rangle_{E_{\text{REF}}} + \langle E_{\text{TARG}} - E_{\text{REF}} \rangle_{E_{\text{TARG}}}) \quad (11)$$

Furthermore, to demonstrate how the performance of the LRA is affected by the difference between the two potentials, we also examined the n -points LRA between the adjacent mapping potentials given by eq 9:

$$\Delta\Delta G_{\text{LRA}}^n = \sum_{m=1}^{n-1} \frac{1}{2} (\langle E_{m+1} - E_m \rangle_{E_m} + \langle E_{m+1} - E_m \rangle_{E_{m+1}}) \quad (12)$$

Obviously, with the large n , this expression converges to the regular FEP.

In the present case, we start our examination of the validity of the LRA treatment by evaluating the FEP and LRA perturbations between the PM3/MM³⁸ and PM6/MM³⁹ potentials of the MOPAC2009⁴⁰ package, with the correct treatment of the solute polarization^{1,7} described in section II.3.3 (see also simulation 1 in the Supporting Information for additional computational details). This is done for an S_N2 reaction in water between methyl chloride and chloride (here PM3/MM is the RP, and PM6/MM is the TP).

For the FEP treatment at the TS, we used (in addition to the QM/MM potential) the harmonic potential:

$$E_{\text{CONS}} = K(\xi - \xi_0)^2 \quad (13)$$

where ξ is the reaction coordinate defined as

$$\xi = R(C - Cl_A) - R(C - Cl_B) \quad (14)$$

Note that $\xi_0(x_{\text{PM3}}^\pm) = \xi_0(x_{\text{PM6}}^\pm) = 0$.

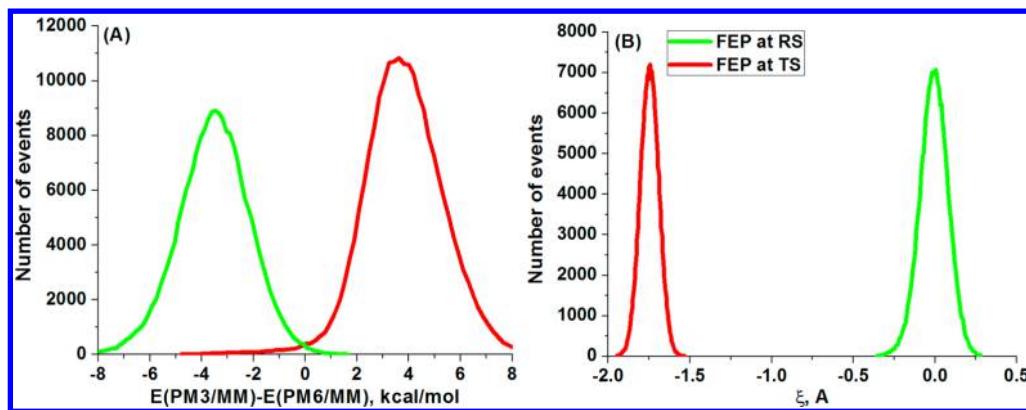


Figure 3. (A) The distribution of the energy gap, $E_{\text{PM3/MM}} - E_{\text{PM6/MM}}$, obtained from the MD trajectories runs used in FEP calculations while moving from $E_{\text{PM3/MM}} + E_{\text{CONS}}(\xi)$ to $E_{\text{PM6/MM}} + E_{\text{CONS}}(\xi)$, where $\xi = d(C - Cl_L) - d(C - Cl_A)$ is the solute RC, in the RS region and at the TS regions (designated by red and green, respectively). (B) The sampling distribution of the nuclear RC, obtained from the MD runs used in FEP calculation of moving from $E_{\text{PM3/MM}} + E_{\text{CONS}}(\xi)$ to $E_{\text{PM6/MM}} + E_{\text{CONS}}(\xi)$ in the RS and TS regions (which are designated by red and green, respectively).

We also evaluated the free energy by the thermodynamic integration method (which is very similar to the regular FEP approach), using

$$\begin{aligned} \Delta\Delta G_{\text{TDI}} &= \int_0^1 \frac{\partial G}{\partial \lambda} d\lambda \\ &= \sum_{m=2}^n \frac{1}{2} (\langle E_{\text{TARG}} - E_{\text{REF}} \rangle_{E_m} \\ &\quad + \langle E_{\text{TARG}} - E_{\text{REF}} \rangle_{E_{m-1}}) \Delta\lambda \end{aligned} \quad (15)$$

The $\Delta\Delta G$ was evaluated for the difference between the potentials E_{REF} and E_{TARG} , which correspond, respectively, to $E_{\text{PM3/MM}}$ and $E_{\text{PM6/MM}}$. These potentials were contained within the TS region by

$$E_{\text{REF}} = E_{\text{PM3/MM}} + 100(\xi - \xi_0|_{\xi_0^{\pm}=0})^2 \quad (16)$$

$$E_{\text{TARG}} = E_{\text{PM6/MM}} + 100(\xi - \xi_0|_{\xi_0^{\pm}=0})^2$$

Similarly, for the RS, we have

$$E_{\text{REF}} = E_{\text{PM3/MM}} + 100(\xi - \xi_0|_{\xi_0^{\text{RS}}=-1.75})^2 \quad (17)$$

$$E_{\text{TARG}} = E_{\text{PM6/MM}} + 100(\xi - \xi_0|_{\xi_0^{\text{RS}}=-1.75})^2$$

The performance of the FEP and LRA calculations is compared in Figure 2, where we see remarkably good performance of the end-point LRA treatment with nearly identical results to those of the full FEP treatment. Note that Figure 2 gives estimates of the free energy difference between the RP and the TP, which were modified by adding the constraining harmonic potentials in order to keep the trajectories at the specific RC values (RS and TS). Note, also, that the harmonic potentials used to obtain the sampling within the RS and within the TS have no bias effect at the RC values (specified by the corresponding ξ_0) at which the FEP was calculated, and a small effect in the vicinity of these points.

Of course, a convincing validation requires us to show that the barrier calculated for the TP using the PD approach is the same as the barrier obtained in a separate PMF calculation. In order to be able to compare the $\Delta\Delta g$ of the PD treatment to the corresponding difference between the full PMF calculations we have to take into account the fact that what we get from the

PMF corresponds to the incomplete partition functions, q 's, of eq 4, whereas the FEP simulations described above sample the whole phase space. Thus, strictly speaking, in order to compare with the difference between the individual PMFs, we have to determine $\Delta\Delta g(\xi)$ through the difference between the free energy functions of the RC (at its specific values with removing the effect of the constraining harmonic potential). This, in turn, can be compared to the $\Delta\Delta G$ estimates obtained by the FEP treatment (eqs 8–15) between the potentials of eqs 16 and 17. This can be done with a treatment, which is a variant of the EVB FEP/US mapping formulation.^{11,12,41} Namely,

$$\begin{aligned} \Delta\Delta g_{\text{PM3/MM} \rightarrow \text{PM6/MM}}(\xi) \\ = \Delta g_{\text{PM6/MM}}(\xi) - \Delta g_{\text{PM3/MM}}(\xi) \end{aligned} \quad (18)$$

where the free energy functions are given by

$$\begin{aligned} \Delta g_{\text{PM3/MM}}(\xi) &= \Delta\Delta G_m \\ &\quad - kT \ln \left\langle \delta(x - \xi) \exp \left(\frac{E_{\text{PM3/MM}} - E_m}{kT} \right) \right\rangle_{E_m} \end{aligned} \quad (19)$$

and

$$\begin{aligned} \Delta g_{\text{PM6/MM}}(\xi) &= \Delta\Delta G_m - kT \\ &\quad \ln \left\langle \delta(x - \xi) \exp \left(\frac{E_{\text{PM6/MM}} - E_m}{kT} \right) \right\rangle_{E_m} \end{aligned}$$

where

$$E_m = (1 - \lambda_m)E_{\text{PM3/MM}} + \lambda_m E_{\text{PM6/MM}} + E_{\text{CONS}}$$

In this expression, we deal with the change from E_{REF} to E_{TARG} (in the EVB treatment, we go from E_1 to E_2) and $\Delta\Delta G_m$ is the result of moving from $E_{\text{PM3/MM}} + E_{\text{CONS}}$ to E_m by the FEP procedure. The second term removes the effect of the constraining potential and selects the particular values of RC.

To demonstrate the point discussed above, we start by building a histogram for the MD trajectories to see the distribution of the energy gap between the two potentials, $(E_{\text{PM3/MM}} - E_{\text{PM6/MM}})$. As can be seen from Figure 3A, the distribution of the energy gap for the MD trajectories propagated during calculation of the free energy of moving

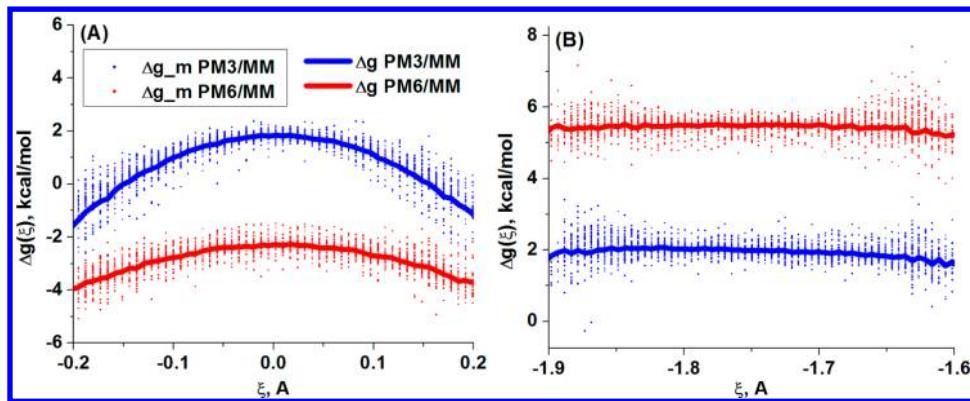


Figure 4. The free energy functions of the solute RC, ξ , obtained at the RC values corresponding to the TS (A) and to the RS (B), for the full mapping between PM3/MM and PM6/MM. The blue line designates $\Delta g_{\text{PM6/MM}}$ and the red line designates $\Delta g_{\text{PM3/MM}}$, obtained as weighted averages over all mapping potentials, blue and red dots, respectively.

from the RP to the TP at the RS is centered at -3.5 kcal/mol. Similarly, the corresponding gap evaluated near the TS is centered at 4 kcal/mol. These are the most probable differences between the two potentials constrained at the RS and at the TS. Adding up these estimates gives an estimate of the total difference between the activation free energy barriers of the RP and of the TP, which is 7.5 kcal/mol. These estimates are close to the values obtained by performing the FEP between the two potentials (see Figure 2A and B).

Further examination of the distribution of the nuclear RC is given in Figure 3B. The figure shows that the harmonic potentials effectively narrowed the sampling to the corresponding target values of the RC. Thus, one can evaluate the free energy perturbations at any value of the RC specified by the ξ_0 parameter of the harmonic potential. While the ξ_0 parameter specifies the RC value at which the free energy perturbation is estimated, the force constant controls the width of the RC distribution. Note, however, that all other degrees of freedom are sampled. We also found that, while increasing the force constant further narrowed the distribution in the RC, it had very little effect on the distribution of the energy gap and almost no effect on the FEP estimate.

At this stage, we estimated the free energy functions of the RC at the RS and at the TS regions, $\Delta g_{\text{PM3/MM}}(\xi)$ and $\Delta g_{\text{PM6/MM}}(\xi)$, using eq 19 (note that this treatment removes the bias potential of eqs 16 and 17). The resulting differences between the PM6/MM and PM3/MM free energy functions, $\Delta\Delta g(\xi)$, are shown in Figure 4, e.g., ~ -4 kcal/mol at the TS ($\xi = 0$) and ~ 3.5 kcal/mol at the RS ($\xi = -1.75$). Thus, we demonstrated that the estimates obtained by the full FEP and by the LRA evaluation of the free energy of moving from the RP to the TP (which are modified by the harmonic potentials to keep the sampling within the RS and the TS regions) are in excellent agreement with the difference between the free energy functions at the corresponding RC values (obtained by removing the bias potential using eq 19).

Finally, we compared in Figure 5 our careful evaluation to the difference between the PMFs calculated for PM3/MM and PM6/MM potentials. Each PMF was obtained (see also simulation 2 in the Supporting Information) by introducing a harmonic constraint at different RC values, ξ_0 , along the reaction path, and combining the simulation windows using the WHAM equations.^{42–44} Inspecting the PMFs, one finds that the difference between the activation free energy barriers is about 7.5 kcal/mol, which is in perfect agreement with the

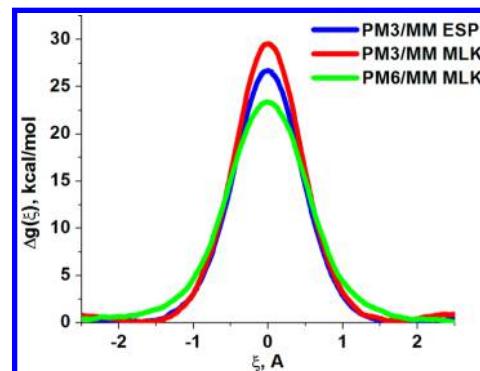


Figure 5. The PMF obtained for the PM3/MM potential with the solute charges derived by the Mulliken analysis (red) and by the ESP-fitting (blue) as well as for PM6/MM potential with the Mulliken solute charges (green). The PMFs have been calculated by the WHAM procedure.

results obtained by the FEP/LRA treatment (eqs 8–15) and by the FEP/US approach (eqs 18 and 19).

In summary, in the above section, we demonstrated by using two MO-type QM/MM potentials that the LRA approach of eq 4 provides a very good approximation for the full FEP/US treatment while giving significant time savings compared to the full FEP treatment. Also note that the LRA approach performs remarkably well when the averages are calculated on both potentials. Finally, by evaluating both the PMFs for the TP and for the RP and the free energy perturbation from the RP to the TP at the TS and the RS, the consistency of our approach is established.

II.2.2. Using the EVB as a Reference Potential. In this section, we will expand on the FEP-LRA evaluation of $\Delta\Delta G$ for moving from the EVB RP to a QM/MM potential. In this case, it is useful to comment here about the evaluation of eq 4 at particular values of the RC, when one uses EVB as the RP. Now, as we have shown in the previous example, the harmonic constraint on the solute RC is quite efficient for estimating the FEP at a particular RC value for an MO-based QM/MM potential (where it is sufficient to have a constraint with $K = 100$ kcal/(\mathring{A}²·mol) to stay at the specified RC). For the EVB potential, we found that a more effective way to keep EVB at its TS (where $E_1 = E_2$) involves using the approximation¹¹

$$E_{\text{EVB}}(x_{\text{EVB}}^{\ddagger}) \approx 0.5E_1 + 0.5E_2 - H_{12} \quad (20)$$

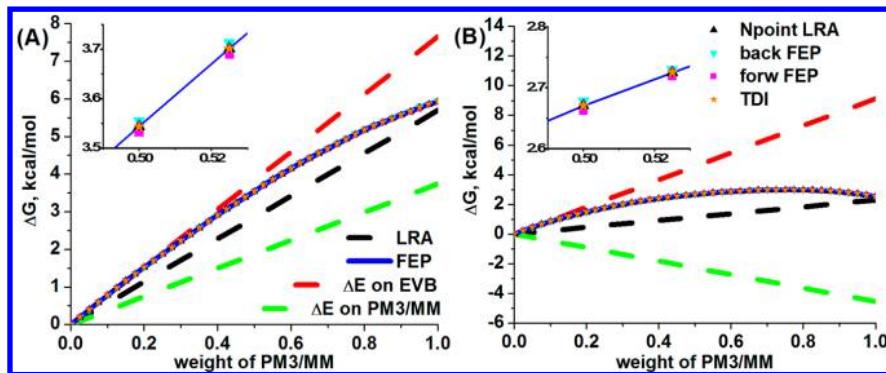


Figure 6. Comparative performance of the LRA approach and of a full multi-step FEP treatment in estimating the free energy of moving from E_{EVB} to $E_{\text{PM3/MM}}$ at the TS (A) and at the RS (B). The estimates were obtained by: 2-point LRA (black dashes); single-point LRA with the average of the energy gap, $E_{\text{PM3/MM}} - E_{\text{EVB}}$, calculated on E_{EVB} (red dashes) and on $E_{\text{PM3/MM}}$ (green dashes); forward FEP (magenta squares); backward FEP (cyan triangles); average FEP (blue line); n -point LRA (black triangles); TDI (orange stars). Note that FEP-based estimates are hardly distinguishable due to overlap (they are shown in insets).

where we use the fact that eq 20 is equal to the adiabatic EVB at the TS:

$$E_{\text{EVB}} = c_1^2 E_1 + c_2^2 E_2 - 2c_1 c_2 H_{12} \quad (21)$$

The use of this expression is equivalent to constraining the eigenvector components to be equal (so that both of the EVB diabatic states equally contribute to the adiabatic state).

Here we estimate the FEP between the EVB and MO-based PM3/MM potentials (see also simulation 3 in the Supporting Information) at the TS using the potentials:

$$\begin{aligned} E_{\text{REF}} &= E_{\text{EVB}}(x_{\text{EVB}}^{\pm}) + 100(\xi - \xi_0|_{\xi_0=0})^2 \quad \text{and} \\ E_{\text{TARG}} &= E_{\text{PM3/MM}} + 100(\xi - \xi_0|_{\xi_0=0})^2 \end{aligned} \quad (22)$$

and at the RS using

$$\begin{aligned} E_{\text{REF}} &= E_{\text{EVB}} + 100(\xi - \xi_0|_{\xi_0=-1.5})^2 \quad \text{and} \\ E_{\text{TARG}} &= E_{\text{PM3/MM}} + 100(\xi - \xi_0|_{\xi_0=-1.5})^2 \end{aligned}$$

The estimate of $\Delta\Delta G$ by the end-point LRA approach was found to provide again a very good approximation to the full FEP (see Figure 6). However, we see the importance of averaging on both the RP as well as on the TP.⁴⁵ The FEPs at the TS and RS are, respectively, ~ 2 and ~ 6 kcal/mol; thus, the activation free energy barrier for EVB is ~ 4 kcal/mol higher than the PM3/MM activation free energy barrier based on this estimate.

In the next step, we constructed (see Figure 7) the histograms of the energy gap $E_{\text{EVB}} - E_{\text{PM3/MM}}$ for the FEP calculations at the RS and at the TS. The histograms show that the most probable difference between the TP and the RP, $E_{\text{PM3/MM}} - E_{\text{EVB}}$, is ~ 7 kcal/mol at the RS and ~ 2 at the TS, which practically coincides with the full FEP estimates and with LRA.

$$\begin{aligned} \Delta g_{\text{PM3/MM}}(\xi) &= \Delta\Delta G_m - kT \\ &\ln \left\langle \delta(x - \xi) \exp \left(\frac{E_{\text{PM3/MM}} - E_m}{kT} \right) \right\rangle_{E_m} \end{aligned} \quad (23)$$

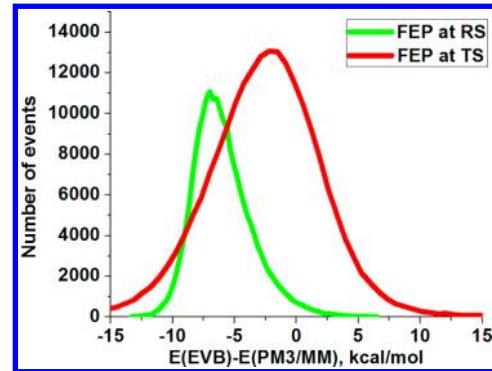


Figure 7. The distribution of the energy gap between the EVB potential and the PM3/MM potential, obtained while performing FEP calculations at the RS (green) and the distribution of the gap between the EVB mapping potential and PM3/MM while doing FEP at the TS (red).

$$\begin{aligned} \Delta g_{\text{EVB}}(\xi) &= \Delta\Delta G_m \\ &- kT \ln \left\langle \delta(x - \xi) \exp \left(\frac{E_{\text{EVB}} - E_m}{kT} \right) \right\rangle_{E_m} \end{aligned}$$

Next, using eq 23, we find that the difference between the free energy functions at the TS, $\xi = 0$, is ~ 4.5 kcal/mol and at the RS ($\xi = -1.5$) is ~ 5.5 kcal/mol (see Figure 8). While this is close to the value found from the FEP calculation at the RS, the difference between the free energy functions at the TS is ~ 3 kcal/mol higher than the corresponding FEP estimate. This is probably due to use of the potential given by eq 20 in the FEP calculation, since if we construct the free energy function along the RC for the approximate EVB potential at the TS (see also Figure 9):

$$\begin{aligned} \Delta g_{\text{EVB}(x_{\text{EVB}}^{\pm})}(\xi) &= \Delta\Delta G_m \\ &- kT \ln \left\langle \delta(x - \xi) \exp \left(\frac{E_{\text{EVB}(x_{\text{EVB}}^{\pm})} - E_m}{kT} \right) \right\rangle_{E_m} \end{aligned} \quad (24)$$

The difference between the free energy functions at the TS is about 3 kcal/mol, which is close to the FEP estimate of 2 kcal/mol. In any case, the correction to the PM3/MM free energy barrier based on eq 23 is about -1 kcal/mol, while on the basis

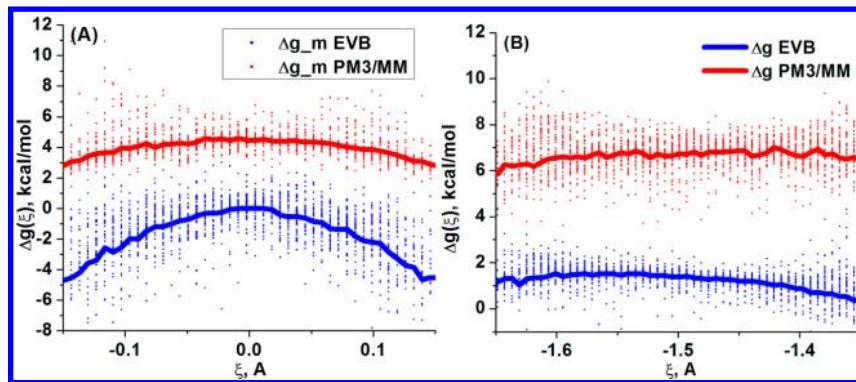


Figure 8. The free energy functions along the reaction coordinate $\xi = d(C - Cl_r) - d(C - Cl_A)$ obtained by FEP calculations of the free energy for moving from the EVB potential to the PM3/MM in water at the TS (A) and at the RS (B). The blue line designates Δg_{EVB} (weight-averaged over the blue points), while the red line designates $\Delta g_{\text{PM3/MM}}$ (weight-averaged over red points).

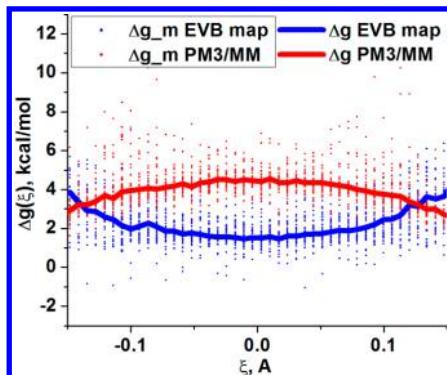


Figure 9. The free energy functions along the reaction coordinate obtained by FEP calculations of the free energy for moving from the EVB mapping potential to PM3/MM in water at the TS. The blue line designates the EVB function (weight-averaged over blue points), while the red line designates the PM3/MM function (weight-averaged over red points).

of the FEP/LRA (with the approximation of EVB potential at the TS by eq 24) it is -4 kcal/mol.

The calculated EVB free energy profile and the PM3/MM PMF are given in Figure 10. The EVB free energy barrier is found to be ~ 1.5 kcal/mol higher than the PM3/MM free energy barrier, and this is in a perfect agreement with the results calculated in the PD model using FEP/US of eq 23. Alternatively, if the EVB free energy profile vs the EVB energy gap is taken as the reference free energy surface, for which the

activation free energy barrier is ~ 3.5 kcal/mol higher than the barrier on the PM3/MM PMF, a good agreement with the corresponding FEP/LRA estimate of 4 kcal/mol is achieved.

The above approach evaluates only the vertical free energy changes, and thus when the TS coordinates on the RP and on the TP are different, we have to evaluate the PMFs on both surfaces at the TS region (see ref 16 for the full discussion of this case). More specifically, if the RC for the QM TS is known (e.g., from the gas phase or from the implicit solvent model minimization), we can evaluate the vertical transition from the QM TS on the RP and then use the PMF (again on the RP) to obtain the free energy of moving from the QM TS on the EVB potential to the actual EVB TS state. In this case, the PMF on EVB can be estimated using the FEP/US,⁶ method as is done in Figure 10. If we do not know the exact QM TS, then PMF on the TP is required. This can be done by one of the approaches described below only in the TS region, which will still give significant time savings compared to the PMF in the full range of the RC. In general, there is no need to use any constraints while performing perturbation at the RS, since the corresponding RSs are effectively sampled when the trajectories are propagated on both the RP and the TP.

II.3. PD Refinement of the Reference Potential by Modifying Its Functional form.

II.3.1. Refining the Intramolecular Potential in the Gas Phase Using Gaussians.

In considering the PD refinement, we noted that it is possible to take advantage of the fact that evaluating the PES in the gas phase for chemical reactions is a relatively trivial operation,

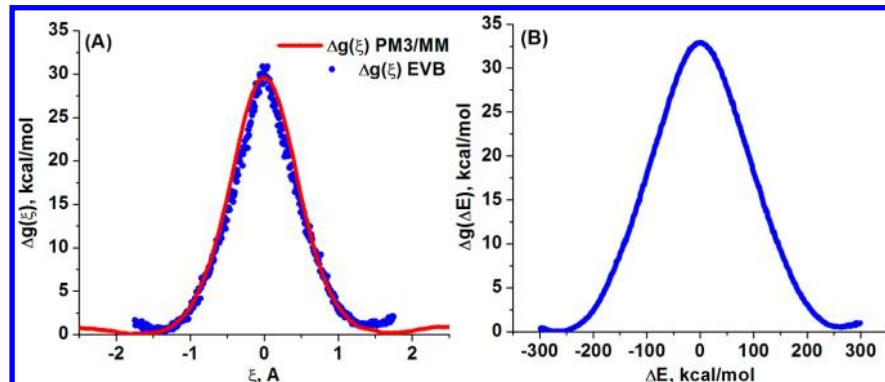


Figure 10. (A) The free energy profile along the nuclear RC obtained by the FEP/US approach using the EVB reference potential (blue dots) and PMF for the PM3/MM target potential (red line); (B) EVB FEP/US free energy profile along the energy gap between the EVB diabatic states.

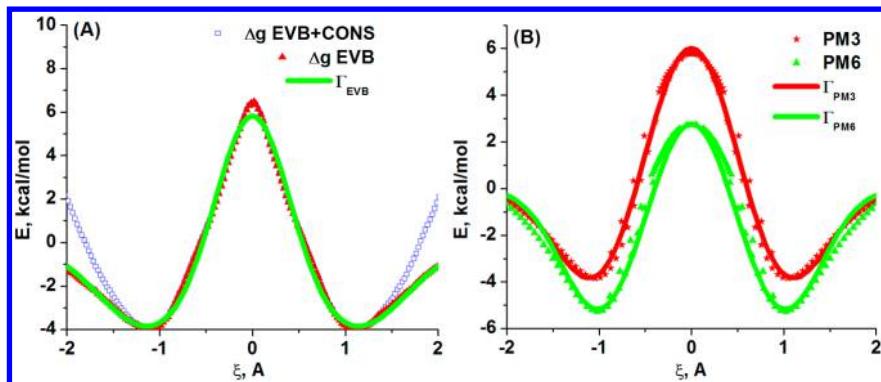


Figure 11. Fitting the Γ -functions. (A) Γ_{EVB} (green line) fitted to the EVB gas-phase PMF/WHAM (red triangles). The EVB potential used in the simulation (blue squares) was modified by removing the effect of the EVB distance constraints for the RC values outside the range $-1 > \xi > 1$. (B) PM3 and PM6 gas phase energy scans from MOPAC2009 (red stars and green triangles) and the corresponding Γ -functions.

which is routinely performed nowadays. This makes PES an easily available source of data for our refinement procedure.

With the gas phase PES, we can simply fit the TP and the RP (e.g., EVB potential) by a set of Gaussian functions. For instance, for the TP, we minimize the least-squares function, using the optimal steepest descent with analytical derivatives, with respect to the parameters of the Gaussians:

$$\mathcal{T}(\{\alpha_k\}, \{A_k\}) = \sum_{i=1}^N (\Gamma_{\text{TARG}}(\mathbf{r}_i) - E_{\text{TARG}}(\mathbf{r}_i))^2 \quad (25)$$

The result of this fitting is a function, Γ , which approximates the TP in the range of the RC where the PES was performed:

$$\Gamma_{\text{TARG}} = \sum_k A_k \exp(-\alpha_k(x - \xi_k)^2) \quad (26)$$

It was found that a combination of 3 Gaussians was sufficient to fit the PES for PM3, PM6, and EVB (see Figure 11 and Table 1). For the details of the fitting algorithm, see also the

Table 1. Parameters for the Γ -Functions Presented in Figure 11 Composed of Three Gaussians: $A_i \exp(-\alpha_i(\xi - \xi_i)^2)$

	A1, A3	A2 (kcal/mol)	α_1, α_3	α_2^{-1}	$\xi_1 (\xi_3)$	ξ_2 (Å)
EVB	-3.92	6.98	1.58	3.66	-1.1 (1.1)	0.0
PM3	-4.39	6.80	2.40	2.22	-1.0 (1.0)	0.0
PM6	-5.25	3.35	2.89	3.70	-1.0 (1.0)	0.0
PM3 cosmo	-0.18	20.85	0.63	2.47	-1.2 (1.2)	0.0
PM6 cosmo	-1.02	16.58	1.34	2.16	-1.1 (1.1)	0.0

Supporting Information. This procedure can be done for both the RP and the TP. In the next step, we use the obtained Γ -functions to refine the original RP (e.g., the EVB RP) using:

$$E'_{\text{EVB}} = E_{\text{EVB}} - \Gamma_{\text{EVB}} + \Gamma_{\text{QM}} \quad (27)$$

This refinement procedure is aimed to bring the RP close to the TP in a given range of the RC, since:

$$E_{\text{TARGET}} \approx E'_{\text{REF}} = E_{\text{REF}} - \Gamma_{\text{REF}} + \Gamma_{\text{TARG}} \quad (28)$$

Now we have the refined RP that provides an improved approximation for the TP which is essential for good convergence of the LRA. This approach is particularly useful when the EVB functional form is sophisticated and the refinement of its parameters is lengthy and tedious.

As a quick demonstration of the described above refinement procedure, we consider an example with an EVB potential serving as the RP and the PM6 potential serving as the TP (see also simulation 4 in the Supporting Information). Figure 12

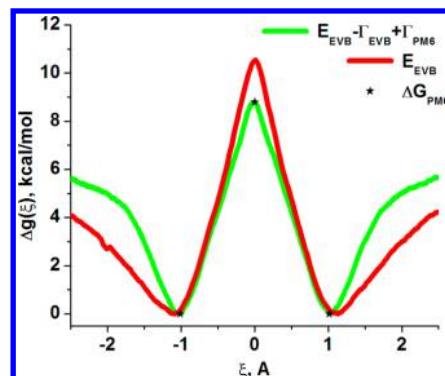


Figure 12. The gas-phase EVB PMF obtained by using the WHAM approach and the harmonic constraints on different values of RC. The red line is for PMF on the original EVB reference potential, while the green line is for the PMF obtained on the refined EVB reference potential ($E_{\text{EVB}} - \Gamma_{\text{EVB}} + \Gamma_{\text{PM6}}$). The black stars are the free energy points calculated on PM6 (target potential) using the harmonic approximation for activation entropy.

illustrates the effect of the Γ -correction refinement (using eq 27) of the original EVB RP on the reference free energy surface. As seen from the figure, the calculated PMF on the original EVB RP is 10.5 kcal/mol, whereas the PMF on the refined EVB RP is 8.8 kcal/mol, which is in excellent agreement with the activation free energy barrier of 8.9 kcal/mol for PM6 (calculated using the harmonic approximation for the entropic contribution). As can be seen, this relatively simple functional form is quite effective in reducing the difference between reference FES and the target FES.

II.3.2. Application of the Gamma-Correction in PMF Calculations: Derivation. Obviously the main application of QM/MM calculations is in studies of reactions in condensed phases. However, prior to extension of the refinement approach proposed in section II.3.1 to the condensed phases, we would like to report several practical applications of this approach to calculating the PMF, which were found to be quite useful and can be also applied to condensed phase PMF calculations.

Since PMF calculations can be quite expensive, it is important to look for ways to optimize the corresponding

calculations. Here we tried to exploit an element of the MTD approach (although, as will be clarified in the concluding discussion, with an entirely different philosophy). That is, as pointed out by van Gunsteren and co-workers³² and as implemented elegantly in the MTD³³ approach, flattening the original potential by iteratively adding to it the negative of the original potential, approximated by a sum of Gaussians, improves the convergence of free energy sampling. Implementation of this strategy in the MTD approach results in a high computational cost of building the negative potential; see the detailed discussion in our recent paper.⁹ However, here we bypassed this cost by using the idea of the PD refinement described in section II.3.1. Namely, we derive the negative potential by fitting the corresponding PES in the gas phase by a sum of Gaussians using eqs 25 and 26. Thus, in addition to their use in the PD refinement, the Γ -functions of eq 26 help in convergence of the PMF calculations. This is achieved by flattening the original potential by adding to it the negative of the corresponding Γ -function. That is, following the FEP/US approach to PMF,³⁴ we construct the mapping potentials of the form

$$E_m = (1 - \lambda_m)E_{\text{RS}} + \lambda_m E_{\text{PS}} \quad (29)$$

where

$$E_{\text{RS}} = E_{\text{QM}} - \Gamma_{\text{QM}} + E_{\text{CONS}}(\xi_{\text{RS}}) \quad (30)$$

$$E_{\text{PS}} = E_{\text{QM}} - \Gamma_{\text{QM}} + E_{\text{CONS}}(\xi_{\text{PS}})$$

where one can, for instance, use E_{CONS} of eq 13.

To obtain the PMF, we start with the FEP approach for the corresponding change in free energy (using eqs 8 and 10). Then, the PMF is evaluated by using a modification of eq 19 (FEP/US approach):

$$\Delta g(\xi) = \Delta G_m - kT \ln \left\langle \delta(x - \xi) \exp \left(-\frac{E_{\text{TARG}}(x) - E_m(x)}{kT} \right) \right\rangle_{E_m} \quad (31)$$

Furthermore, the results from different simulation windows are combined by

$$\Delta \bar{g}(\xi) = \sum_{\text{frames}} \frac{N_i(\xi)}{\sum_{\text{frames}} N_i(\xi)} \Delta g_i(\xi) \quad (32)$$

where $N_i(\xi)$ is the number of times MD visited a particular RC value, ξ , while propagating on the i th mapping potential. The practical application of this approach to PMF calculations, when the solvent is modeled using an implicit solvent model, will be illustrated below.

II.3.3. Application of the Gamma-Correction in PMF Calculations: Practical Examples for Condensed Phase Calculations. When moving to a condensed phase, we should take into account the fact that different QM(ai)/MM potentials will have different charge distributions for the solute, and that this has to be reflected in the RP. A straightforward application of the Γ -function approach (described in section II.3.2) to condensed phases is to evaluate the PES with a solvent model, and subsequently derive the corresponding Γ -function that reflects the solvation effects. This strategy is obviously impractical when one deals with explicit solvent models, which are extremely expensive for QM(ai)/MM studies.

However, performing reasonable minimization is not expensive when one uses an implicit solvent model.

To illustrate this approach, as a first step, we examined the option of using the Γ -function which captures the solvation effect by the COSMO model.⁴⁶ As a test system, we took the same S_N2 reaction described above and fitted the Γ_{COSMO} function to the PES obtained using MOPAC2009⁴⁰ with PM3 + COSMO combination (see Figure 13). The free energy

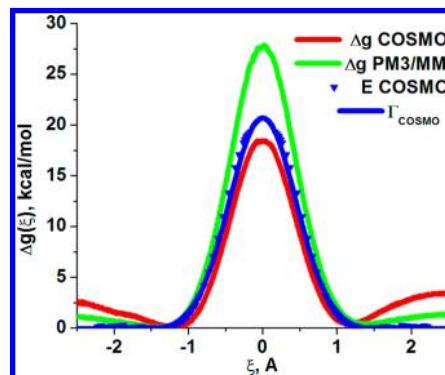


Figure 13. The PMFs obtained for the PM3 potential with explicit (green) and implicit COSMO (red) solvation models. The PMFs were obtained by adding to the PM3(COSMO) and to the PM3/MM potentials the negative of Γ_{COSMO} . The figure also depicts the Γ_{COSMO} (blue line) function, derived from PES performed on the PM3 potential with the COSMO solvation model (blue triangles).

profile, $\Delta g_{\text{PM3,COSMO}}$, was obtained by the approach described in section II.3.2 with the MD simulation run on the flat potential, $E_{\text{PM3,COSMO}} - \Gamma_{\text{COSMO}}$. The high efficiency of sampling at the TS region and its uniform distribution along the studied range of the RC is shown in Figure 14.

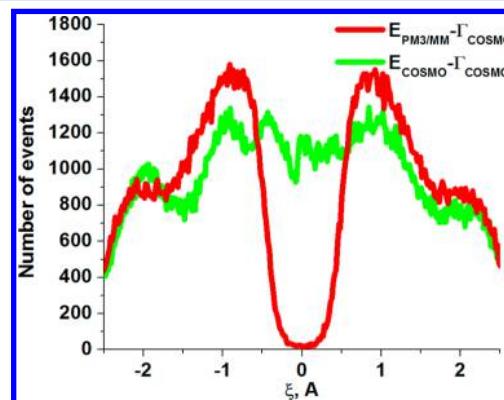


Figure 14. The distribution of the nuclear RC during MD simulation on the potential $E_{\text{PM3/MM}} - \Gamma_{\text{COSMO}}$ (red) and on the potential $E_{\text{PM3+COSMO}} - \Gamma_{\text{COSMO}}$ (green). The histogram for $E_{\text{PM3/MM}} - \Gamma_{\text{COSMO}}$ shows a poor sampling at the TS for the explicit model, since Γ_{COSMO} does not make the explicit potential fully flat.

The proposed approach provides an effective and powerful way that outperforms the recent attempts to obtain the FES of solute in implicit solvent models.⁴⁷ In particular, the FEP/US PMF allows movement along the specified reaction path in the most efficient way as far as the sampling is concerned (since the FEP increments are chosen in terms of the energy difference between the two adjacent states which is what essentially matters the most for the fastest free energy convergence). The

representative Γ -function flattens the original potential and within a given simulation window, and thus, it naturally allows for proper sampling in all directions perpendicular to the RC.

Unfortunately, the implicit solvent models have difficulties in capturing changes of the solute cavity during chemical reaction (even with a reasonable calibration) and can miss other microscopic features of solvent. Thus, we may have significant errors in estimates of the TS solvation (see Table 2 and simulation 5 in the Supporting Information). Furthermore, the implicit models cannot capture the microscopic physics of protein interiors.

Table 2. Comparing the Solvation Energies for the COSMO and for the Explicit SCAAS Models

	$\Delta G_{AC}(0 \rightarrow Q_{\text{gas}})$, kcal/mol	$\Delta H_{\text{f,COSMO}} - \Delta H_{\text{f,GAS}}$, kcal/mol
X_{TS} PM3 (fixed)	-52.1	-61.1
X_{RS} PM3 (fixed)	-67.8	-70.4
$\Delta\Delta G_{\text{solv}}$	15.7	9.3

In an attempt to overcome the limitations of implicit solvent models, we considered the performance of the proposed approach for the explicit solvent models. We obtained the PMF with the explicit solvent model, $\Delta g_{\text{PM3/MM}}$ (using the ESP-derived QM charges) by running on the potential

$$E_m = \langle \Psi_{\text{polar}} | H'_{\text{PM3/MM}} | \Psi_{\text{polar}} \rangle - \Gamma_{\text{PM3,COSMO}} + (1 - \lambda_m)K(\xi - \xi_{\text{RS}})^2 + \lambda_m K(\xi - \xi_{\text{PS}})^2 \quad (33)$$

with the polarized PM3/MM Hamiltonian

$$h'_{\text{PM3/MM}}(i) = h_{\text{PM3}}(i) + \sum_{\text{MM atoms}} \frac{q_j}{r_{ij}} \quad (34)$$

which we have implemented in MOPAC2009, following our earlier work⁷ (see also simulation 6 in the Supporting Information). While the resulting PMF (see Figure 13) is in excellent agreement with the PMF separately obtained by the WHAM approach (see simulation 2 in the Supporting Information), the efficiency of sampling on the potential given by eq 33 at the TS (shown in Figure 14) is very low. This can be explained by comparing the barriers in the explicit and the implicit solvation models. The addition of $-\Gamma_{\text{COSMO}}$ does not make the TS region on the explicit potential fully flat due to the differences in the TS solvation shown in Table 2.

The PMF on the PM3 + COSMO potential resulted in a lower free energy barrier than that obtained by the energy scan. Furthermore, with the explicit solvent model, we also obtained a higher free energy barrier. To explain the difference in the activation free energy between different solvent models, we calculated the solvation free energies for the gas-phase optimized PM3 RS and TS using MOPAC2009 with the COSMO⁴⁶ model and by the adiabatic charging model.³⁵ From the results given in Table 2, one can see that there is a difference in solvation of the RS and of the TS between the explicit and implicit models, with the highest disagreement between the implicit and explicit solvent models found for the TS solvation. Essentially for this S_N2 reaction, $\Delta\Delta G_{\text{SOLV}}$ is a rough approximation of the contribution to the activation free energy barrier due to the solvation difference (if we add the gas phase barrier of ~ 12 kcal/mol, we get roughly the PM3/MM barrier given in Figure 13). The deficiencies of the barrier

obtained by this specific COSMO implementation are beyond the scope of this work, since different implicit models can give different results and it is always important to compare the solvation estimates by the different models (see, e.g., the careful study of ref 48). In general, any attempts to obtain a reliable QM/MM surface should involve careful calibration of the model on the observed solvation energies (as we have been doing with the EVB model for a very long time), but this is not the issue addressed in the present work.

In section II.3.3, we showed that the approach of section II.3.2 provides an extremely powerful way of evaluating the PMF with an implicit solvent model. This is significant, since the implicit solvent models are frequently used in studies of reference solution reactions, and such studies require a very tedious manual mapping in order to obtain reliable estimates.⁴⁹ Now we have provided a systematic and effective way of obtaining the relevant free energy surfaces.

Mapping with an implicit solvent model does not present major problems (except missing the nonequilibrium solvation effect⁴⁵ which would lead to underestimating the implicit solvation barrier) as well as having other well-known disadvantages of continuum models. However, in section II.3.3, we also saw that evaluation of the PMF becomes much more challenging when we deal with explicit solvent models. Here the selection of the proper mapping potential is not trivial, as we have to force the combined solute–solvent coordinate to respond to the change in the solute charge distributions. Here we can take advantage of the fact that any RP-based QM/MM calculation in polar environments (water or enzymes), and, in particular, in the approaches that use semiempirical MO-based QM/MM potentials as the RP,^{19,25} can be refined using the Γ -correction of eq 28. This correction can be derived, for instance, by eqs 25 and 26 with an implicit solvent model for the original RP and for the TP (e.g., COSMO for which the analytical gradients are available), using:

$$E'_{\text{REF}} = E_{\text{REF}} - \Gamma_{\text{REF,IMPL}} + \Gamma_{\text{TARG,IMPL}} \quad (35)$$

Even though evaluating the PES with the implicit solvent is relatively expensive, it describes more consistently the solute changes along the reaction path in the condensed phases than the corresponding gas phase PES, since the solute polarization is captured in a physically more reasonable way by the implicit solvent models.

In view of the length of section II.3, we summarize below what was done in this section. Section II.3.1 showed how to refine the RP using the Γ -correction. Section II.3.2 introduced another practical use of the Γ -correction for improving the sampling efficiency while evaluating PMF. Section II.3.3 demonstrated a successful straightforward application with an implicit solvent model and less efficient (but possible if agreement between the implicit and explicit solvation models is improved) with the explicit solvent. Moreover, eq 35 provides a practical recipe for refinement of RP in the condensed phase described in section II.3.1.

II.4. Extension to the Condensed Phase Using the EVB-Type Solvent Potential. In this section, we review application of the methods described in section II.3 when they are combined with an idea of using the EVB-type solvent driving potential. This allows for implementation of eq 30 in a practical way when the difference between the implicit and explicit models is significant or when it is desirable to have transferable Γ -functions (e.g., from the gas phase to water or from water to protein). These functions can be easily derived in

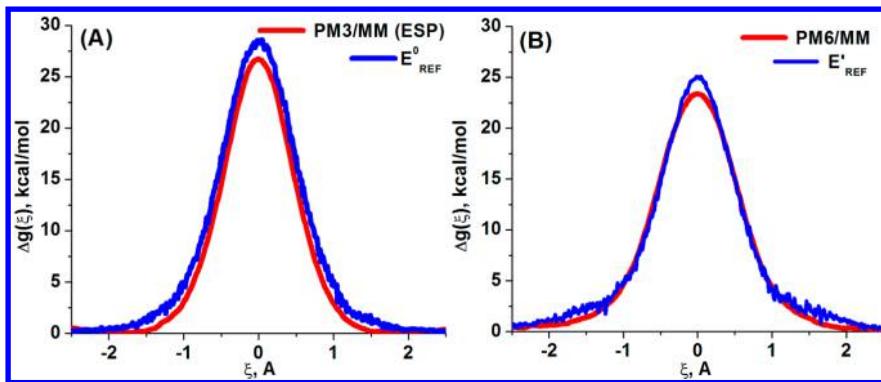


Figure 15. The PMFs obtained for the CH_3Cl and $\text{Cl}^- \text{S}_{\text{N}}2$ reaction in water. (A) (red) PMF constructed for the $E_{\text{PM3/MM}}$ potential (polarized PM3 Hamiltonian with the solvent polarized by the ESP charges) using the WHAM approach. (blue) PMF obtained by sampling on the flat $E_{\text{PM3(GAS)}} - \Gamma_{\text{PM3(GAS)}}$ with the EVB-type solvent driving potential using the FEP/US approach for the potential $E_{\text{PM3(GAS)}} + E_{\text{ss}}(Q_{\text{PM3}})$. (B) Demonstrating the refinement of the original reference potential, $E_{\text{PM3(GAS)}} + E_{\text{ss}}(Q_{\text{PM3}})$, by forcing it to approximate the PM6/MM target potential. (red) PMF for the $E_{\text{PM6/MM}}$ potential (polarized PM6 Hamiltonian with the solvent polarized by the ESP charges) constructed using WHAM. (blue) The refined reference free energy surface obtained by sampling on the flat $E_{\text{PM3(GAS)}} - \Gamma_{\text{PM3(GAS)}}$ with the EVB-type solvent driving potential, using the FEP/US approach, for $E_{\text{TARG}} = E_{\text{PM3}} - \Gamma_{\text{PM3}} + \Gamma_{\text{PM6}} + E_{\text{SOLV}}(Q_{\text{PM3}})$.

the gas phase (or with an implicit solvent model) and subsequently applied to calculations in condensed phases without reparameterization. The Γ -correction, fitted to the gas phase calculations, contains only information about changes in the solute along the reaction path, namely, the information about the intramolecular part of the potential. To incorporate the effect of solvent and still take advantage of the gas-phase derived Γ -function, we consider a mapping potential where the QM and MM regions are coupled classically by

$$E_{\text{map}} = E_{\text{PM3,gas}} - \Gamma_{\text{PM3,gas}} + (1 - \lambda)(E_{\text{CONS}}(\xi_{\text{RS}}) + E_{\text{ss}}(\mathbf{Q}_{\text{RS}})) + \lambda(E_{\text{CONS}}(\xi_{\text{PS}}) + E_{\text{ss}}(\mathbf{Q}_{\text{PS}})) \quad (36)$$

where \mathbf{Q}_{RS} and \mathbf{Q}_{PS} are the vectors of the QM solvated charges at the RS and at the PS.

$$E_{\text{ss}}(\mathbf{Q}_k) = \sum_{i=1}^{\text{QM}} \sum_{j=1}^{\text{MM}} \left(\frac{Q_i^{(k)} q_j}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \quad (37)$$

Equation 37 contains the EVB-type solvent driving potential, which polarizes the solvent in the correct direction toward the product state by the solvated solute charges, and captures the nonequilibrium solvent effect. However, this treatment should be applied with care to $\text{S}_{\text{N}}1$ and other charge separation reactions¹¹ (since essentially the QM region is described by the gas phase QM potential, which is perturbed by the MM force field without the correct solute polarization). This mapping potential allows separating the QM and MM parts for the straightforward application of the gas-phase Γ -functions, while addressing the sampling problems separately (namely, requiring that the intra QM barrier is flattened with the gas-phase Γ -function and the solvent is pulled toward the TS by the EVB-type potential).

A correct treatment of the solute polarization in the case of a MO-based QM/MM potential is obtained by:

$$E_{\text{QM/MM}} = \langle \Psi_{\text{polar}} | H'_{\text{QM/MM}} | \Psi_{\text{polar}} \rangle + E_{\text{Vdw}} \quad (38)$$

Note that this treatment was used for most of the calculations in this work.

At this stage, it is interesting to see whether we can extend the approach of eq 28 to simulations of processes in polar

environments. As was mentioned above, the gas phase Γ -function only contains information about the intramolecular contributions and says nothing about the intermolecular interactions with solvent. Equation 35 with the Γ -function obtained with an implicit solvent can provide a higher-level approximation, but this requires further examination. In particular, it is interesting to examine whether such a correction (obtained by fitting the PES performed with an implicit solvent model) can be directly used to refine the RP for enzymatic simulation. Similarly, application of the EVB-type solvent driving potential combined with eq 33 should be further explored for purposes of improving the sampling of a MO-based QM/MM RP including the correct solute polarization.

A hypothetical situation of refining a general RP (e.g., a MO-type QM/MM PM3/MM potential) is considered in Figure 15A. Here the original reference potential is given by

$$E_{\text{REF}} = E_{\text{PM3,gas}} + E_{\text{ss}}(\mathbf{Q}_{\text{PM3,gas}}) \quad (39)$$

(with the corresponding FES obtained after removing the bias of eq 36 using eq 31, and combining the multiple simulation windows using eq 32) (see simulation 7 in the Supporting Information). The corresponding original FES (as well as the PMF obtained using eq 38) are given in Figure 15A. To refine this reference potential for the PD calculation of the free energy barrier on the PM6/MM potential, we use

$$E'_{\text{REF}} = E_{\text{REF}} - \Gamma_{\text{PM3,gas}} + \Gamma_{\text{PM6,gas}} \quad (40)$$

with the mapping potential given by eq 36. The refined reference FES (for the potential of eq 40) as well as the target PM6/MM FES (obtained using the PM6/MM version of eq 38) are given in Figure 15B. In actual QM(ai)/MM studies, one can substitute the PM6/MM TP (which was used here only for demonstration purposes) by an expensive QM(ai)/MM potential. In such cases, it is reasonable to assume that the corresponding gas phase potential and the Γ -function of eq 40 can be substituted with the corresponding correction obtained with an implicit solvent model (e.g., COSMO).

Turning now to the EVB RP, we note that this model describes the correct polarization of solute by including the effect of solvent in the EVB diagonal elements.¹¹ Thus, for the EVB RP, the Γ -correction in the gas phase is the correction for

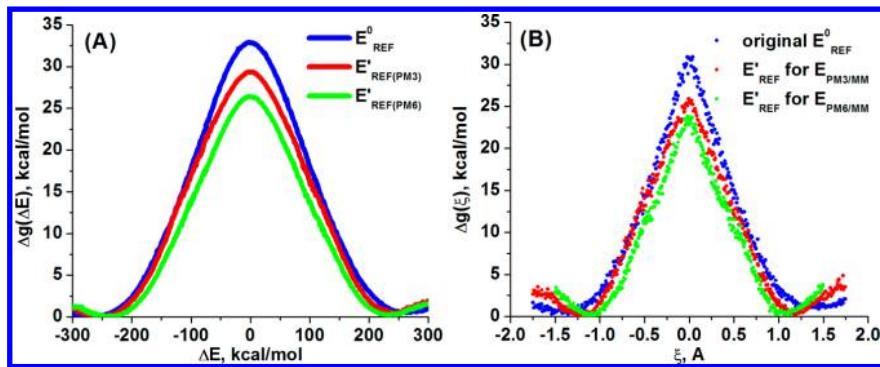


Figure 16. Demonstrating the refinement of the EVB reference potential in condensed phases, using the correction potential $-\Gamma_{\text{EVB}} + \Gamma_{\text{TARG}}$ fitted by Gaussians plus a vector of the new EVB charges derived for the target potential. The free energy profiles obtained by the EVB FEP/US approach along the EVB energy gap (A) and along the nuclear RC (B). (blue) the original reference potential with parameters refined for the gas phase PM3; (red) the refined EVB reference potential for PM3/MM target potential; (green) the refined EVB reference potential for PM6/MM target potential.

the intramolecular interactions within the solute. Moreover, the solute–solvent interaction term can be refined by using

$$E'_{\text{EVB}} = E_{\text{INTRA}} + \Delta\Gamma_{\text{INTRA}} + c_1^2 E_{1,ss}(\mathbf{Q}_{\text{RS}}^{\text{TARG}}) + c_2^2 E_{2,ss}(\mathbf{Q}_{\text{PS}}^{\text{TARG}}) \quad (41)$$

In other words, in the case of the EVB RP, the gas phase correction, $\Delta\Gamma$, refines the intramolecular part of the EVB RP by the relationship:

$$\Delta\Gamma_{\text{INTRA}} = \Gamma_{\text{TARG}} - \Gamma_{\text{EVB}} \quad (42)$$

Since the original EVB RP can be partitioned as

$$E_{\text{EVB}} = c_1^2 E_1 + c_2^2 E_2 - 2c_1 c_2 H_{12} = E_{\text{INTRA}} + c_1^2 E_{1,ss}(\mathbf{Q}_{\text{RS}}^0) + c_2^2 E_{2,ss}(\mathbf{Q}_{\text{PS}}^0) \quad (43)$$

The only minor problem here can be that the intramolecular interactions reflected by the Γ_{EVB} are calculated with the original vector of EVB charges, \mathbf{Q}^0 , and one might need to recalculate the Γ_{EVB} for the new set of charges to improve the accuracy of the $\Delta\Gamma_{\text{INTRA}}$ correction.

In conclusion, we demonstrate in Figure 16 how we refine the EVB RP originally derived for the PM3 gas-phase calculations. In other words, we refine the EVB RP for PM3/MM and PM6/MM as the TPs, see also simulation 8 in the Supporting Information.

III. CONCLUDING DISCUSSION

This work refines, quantifies, and validates our paradynamics model. The validation starts by comparing the LRA approach to the full FEP treatment for two arbitrary MO-type QM/MM potentials, and is also carried out for EVB as a RP. The LRA is found to be sufficiently accurate in evaluation of the free energy perturbation when moving from the RP to the TP at the TS and at the RS. It is also found that it is sufficient to use a harmonic constraint potential in determining the free energy correction at the specific value of the RC. Overall, it is concluded that the calculated $\Delta\Delta G$ between the two modified potentials is a good approximation for the difference between the free energy functions, $\Delta\Delta g(\xi)$, of the TP and of the RP at the coordinate ξ . That is, it is found that

$$\begin{aligned} \Delta\Delta G_{\text{LRA}}(E_{\text{REF}} + E_{\text{CONS}}(\xi) \rightarrow E_{\text{TARG}} + E_{\text{CONS}}(\xi)) \\ \cong \Delta\Delta G_{\text{FEP}} \\ \cong \Delta g_{\text{TARG}}(\xi) - \Delta g_{\text{REF}}(\xi) \end{aligned} \quad (44)$$

Several practical improvements for the PD reference potential are proposed. More specifically, a novel idea of refining the RP (usually following the initial refinement of the EVB parameters in the case of the EVB RP) in a general case is put forward and tested on the $\text{CH}_3\text{Cl} + \text{Cl}^- \text{S}_{\text{N}}2$ reaction in water. The refinement approach is based on modifying the RP with the help of Gaussian functions fitted to the PESs along the RC performed on the original RP and on the TP. This technique is shown to be efficient for gas phase simulations, as well as with the implicit solvent models with no further modifications. Moreover, the extension of this model to simulation in condensed phases with the explicit solvent models is made in a reasonable way, although further validation is needed.

It should be clarified here that our approach for adding Gaussian functions shares common features with the MTD basic idea of flattening the original potential by adding to it the negative of the model potential.^{32,33} However, in contrast to spending an enormous amount of time while iteratively elevating the local minima of the ab initio QM/MM potential,⁵⁰ we propose a simple approach of building the negative potential by simply evaluating the gas phase or the implicit solvent PES along the RC.

We note here that some RP approaches, which are very similar to the PD approach (e.g., ref 51), expressed the optimistic perspective that the results obtained therein are superior to our earlier RP results. What has been overlooked is that those works addressed the rather trivial challenge where the solute is fixed (this case has been already handled in our 1992 work⁷). That is, with a fixed solute, the calculations converge very fast. Nevertheless, obtaining reliable solvation free energies by ab initio calculation is still a challenge even with a fixed solute.³ Here we must point out that even a very recent innovative idea of combining lambda dynamics and metadynamics⁵² in solvation calculations is significantly less efficient than our RP approach, where almost all the charging processes are done with a classical potential.⁵³

The present work demonstrated that the end point LRA approach, involving averaging both on the RP and on the TP, practically coincides with the full *n*-step FEP, whereas the averages taken only on the RP result in a higher error. This

point is important, since most of the works^{24,25,51,54} that adopted our RP idea have not yet moved to the end points LRA treatment. Of course, the error associated with the initial LRA estimate will further decrease as the RP is refined to increase the overlap with the TP. In any case, both elements of our PD approach, the refinement of the RP and the LRA estimate of the FEP, are crucial for the efficiency and accuracy when the RP is used. In fact, using the refined PD RP, one can sufficiently sample the TP to ensure the free energy convergence (since the computationally cheap refined RP has high overlap with the TP) as well as to use the highest possible level of theoretical treatment for the QM region given limited computational resources (with evaluation of the true FEP). In fact, we demonstrated in our previous work⁹ that the PD approach is about 2 orders of magnitude more efficient (in terms of calling the ab initio QM calculations) than MTD. These advantages of the PD approach might help in avoiding the possible corresponding artifacts encountered by some researchers in MTD QM/MM studies, which resulted in difficulties of reproducing the experimental catalytic activity of enzymes.⁵⁵ In this respect, we would like to clarify that even the extreme efforts and technically impressive (by the size of the QM region) results in studies of B₁₂ enzymes⁵⁶ by MTD are not necessarily as reliable as one might tend to think considering the elegance of the approach. In fact, the conclusions⁵⁶ that the surface cannot be concerted are problematic, as the actual surface is quite flat in the diagonal range and the actual calculations involve a very short sampling time for the PMF calculations (1.5 ps for each simulation window) and a very short (5 fs) time interval between the iterative depositions of Gaussians in the MTD sampling, which might not be sufficient for the accurate elevation process. Furthermore, we would like to clarify that there is no single experiment that actually excludes the concerted mechanism (clearly no such experiment exists in solution where there is no relevant model system with both the Co–C bond and the hydrogen donor). Here we would like to state that the strategy (e.g., ref 57) of using very careful QM calculations in solution (with a higher level QM model than that used in ref 56) has produced a concerted path, and that moving the solution surface to the protein environment by a calibrated EVB model is expected to be more reliable (both in terms of the sampling and in terms of extrapolation of a reliable reference system) than the direct MTD in the protein site. For more solid conclusions, it is crucial to perform the MTD simulations also for the solution reaction (this is also crucial for obtaining the catalytic effect⁵⁷). Furthermore, studies of enzyme catalysis require major experience and validation in treating electrostatic effects in proteins and such validations have not yet been reported by the MTD studies, whereas this has been done reliably by reference potential QM(ai)/MM calculations of pK_a's.³ Here again, it would be interesting to see a comparison between the PD and MTD studies, and since the EVB (with a very careful calibration) has already reproduced quantitatively the catalysis in B₁₂ with a concerted path, we tend to believe that the same results (or at least a flat surface) will be obtained by the PD QM(ai)/MM approach with the EVB RP.

As clarified in our previous work, the MTD is formally similar to the earlier PD approach⁹ in the sense that the RP is built iteratively. However, the important difference, which results in major time savings in the PD approach, is that, while the search of the negative RP is carried out blindly by directly sampling of the TP in MTD calculations, PD calculations start

with a RP that is already close to the TP and further refining it iteratively. Of course, our PD approach requires an approximate knowledge of the reaction path, but as we pointed out in detail elsewhere,⁹ we find it useful not to have a black box blind approach in studies of chemical problems. Our recipe of overcoming the blind search is quite simple. It involves the initial mapping with an implicit solvent model that can provide the main mechanistic options, the relevant RC, and the reaction paths. It is true that there is a conceivable advantage in fully automated black box approaches (provided they are sufficiently efficient). However, when one faces the dilemma, in principle, between the simplified “artificial intelligence” and human intellect and experience then, knowing the fundamental chemical concepts, or identifying the possible reaction path in solution using fast models, the latter has an advantage in solving complex problems. Thus, we consider the present work as another illustration of the advantage and power of the PD in QM/MM studies.

■ ASSOCIATED CONTENT

Supporting Information

Computational details of the performed simulations and description of the implemented MOPAC-MOLARIS QM/MM interface and the script in Maple9.5 used in fitting potential energy scans with sums of Gaussians. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: warshel@usc.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by NIH grant GM 24492 and NSF grant MCB-08364000. We would like to express our profound gratitude to Dr. J. J. Stewart for letting us work with the code of MOPAC 2009 in these QM/MM studies. We gratefully acknowledge the University of Southern California's High Performance Computing and Communications Center for computer time.

■ REFERENCES

- (1) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227–249.
- (2) Hu, H.; Yang, W. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601.
- (3) Kamerlin, S. C. L.; Haranczyk, M.; Warshel, A. *J. Phys. Chem. B* **2008**, *113*, 1253–1272.
- (4) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
- (5) Ranaghan, K. E.; Mulholland, A. *J. Int. Rev. Phys. Chem.* **2010**, *29*, 65–133.
- (6) Muller, R. P.; Warshel, A. *J. Phys. Chem.* **1995**, *99*, 17516.
- (7) Luzhkov, V.; Warshel, A. *J. Comput. Chem.* **1992**, *13*, 199–213.
- (8) Bentzien, J.; Muller, R. P.; Florian, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 2293–2301.
- (9) Plotnikov, N. V.; Kamerlin, S. C. L.; Warshel, A. *J. Phys. Chem. B* **2011**, *115*, 7950.
- (10) Warshel, A.; Weiss, R. M. *J. Am. Chem. Soc.* **1980**, *102*, 6218–6226.
- (11) Hwang, J. K.; King, G.; Creighton, S.; Warshel, A. *J. Am. Chem. Soc.* **1988**, *110*, 5297–5311.
- (12) King, G.; Warshel, A. *J. Chem. Phys.* **1990**, *93*, 8682.
- (13) Zwanzig, R. *J. Chem. Phys.* **1954**, *22*, 1420.
- (14) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.

(15) Lee, F. S.; Chu, Z.-T.; Bolger, M. B.; Warshel, A. *Protein Eng.* **1992**, *5*, 215–228.

(16) Rosta, E.; Klahn, M.; Warshel, A. *J. Phys. Chem. B* **2006**, *110*, 2934.

(17) Mones, L.; Kulhanek, P.; Simon, I.; Laio, A.; Fuxreiter, M. *J. Phys. Chem. B* **2009**, *113*, 7867–7873.

(18) Claeysens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schütz, M.; Thiel, S.; Thiel, W.; Werner, H.-J. *Angew. Chem., Int. Ed.* **2006**, *45*, 6856–6859.

(19) Lonsdale, R.; Hoyle, S.; Grey, D. T.; Ridder, L.; Mulholland, A. *J. Biochemistry* **2012**, *51*, 1774–1786.

(20) Valiev, M.; Yang, J.; Adams, J. A.; Taylor, S. S.; Weare, J. H. *J. Phys. Chem. B* **2007**, *111*, 13455–13464.

(21) Wang, T.; Yin, H.; Wang, D.; Valiev, M. *J. Phys. Chem. A* **2012**, *116*, 2371–2376.

(22) Woods, C. J.; Manby, F. R.; Mulholland, A. J. *J. Chem. Phys.* **2008**, *128*, 014109–014108.

(23) Wood, R. H.; Yezdimer, E. M.; Sakane, S.; Barriocanal, J. A.; Doren, D. J. *J. Chem. Phys.* **1999**, *110*, 1329–1337.

(24) Beierlein, F. R.; Michel, J.; Essex, J. W. *J. Phys. Chem. B* **2011**, *115*, 4911–4926.

(25) Cao, J.; Bjornsson, R.; Bühl, M.; Thiel, W.; van Mourik, T. *Chem.—Eur. J.* **2012**, *18*, 184–195.

(26) Zhang, Y.; Liu, H.; Yang, W. *J. Chem. Phys.* **2000**, *112*, 3483–3492.

(27) Valiev, M.; Garrett, B. C.; Tsai, M.-K.; Kowalski, K.; Kathmann, S. M.; Schenter, G. K.; Dupuis, M. *J. Chem. Phys.* **2007**, *127*, 051102–051104.

(28) Yin, H.; Wang, D.; Valiev, M. *J. Phys. Chem. A* **2011**, *115*, 12047–12052.

(29) Savelyev, A.; Papoian, G. A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 20340–20345.

(30) Piela, L.; Kostrowicki, J.; Scheraga, H. A. *J. Phys. Chem.* **1989**, *93*, 3339–3346.

(31) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 154–163.

(32) Huber, T.; Torda, A. E.; Gunsteren, W. F. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695–708.

(33) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562–12566.

(34) Kato, M.; Warshel, A. *J. Phys. Chem. B* **2005**, *109*, 19516–19522.

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

(36) Warshel, A.; W. Parson, W. Q. *Rev. Biophys.* **2001**, *34*, 563–679.

(37) Kamerlin, S. C. L.; Warshel, A. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 30–45.

(38) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209–220.

(39) Stewart, J. J. *J. Mol. Model.* **2007**, *13*, 1173–1213.

(40) Stewart, J. J. P. *MOPAC2009* ed.; Stewart Computational Chemistry: Colorado Springs, CO, 2008.

(41) Aqvist, J.; Hansson, T. *J. Phys. Chem.* **1996**, *100*, 9512–9521.

(42) Grossfield, A. *WHAM 2.0.6* ed.; University of Rochester: Rochester, NY, 2012.

(43) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(44) Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275–282.

(45) Villa, J.; Warshel, A. *J. Phys. Chem. B* **2001**, *105*, 7887–7907.

(46) Klamt, A.; Schuurmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

(47) Branduardi, D.; De Vivo, M.; Rega, N.; Barone, V.; Cavalli, A. *J. Chem. Theory Comput.* **2012**, *7*, 539–543.

(48) Kamerlin, S. C. L.; Warshel, A. *J. Phys. Chem. B* **2009**, *113*, 15692–15698.

(49) Sharma, P. K.; Xiang, Y.; Kato, M.; Warshel, A. *Biochemistry* **2005**, *44*, 11307–11314.

(50) Barducci, A.; Bonomi, M.; Parrinello, M. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 826–843.

(51) Rod, T. H.; Ryde, U. *Phys. Rev. Lett.* **2005**, *94*, 138302.

(52) Wu, P.; Hu, X.; Yang, W. *J. Phys. Chem. Lett.* **2011**, *2*, 2099–2103.

(53) Rosta, E.; Haranczyk, M.; Chu, Z. T.; Warshel, A. *J. Phys. Chem. B* **2008**, *112*, 5680–5692.

(54) Retegan, M.; Martins-Costa, M.; Ruiz-Lopez, M. F. *J. Chem. Phys.* **2010**, *133*, 064103–064105.

(55) Stanton, C. L.; Kuo, I. F. W.; Mundy, C. J.; Laino, T.; Houk, K. N. *J. Phys. Chem. B* **2007**, *111*, 12573–12581.

(56) Bucher, D.; Sandala, G. M.; Durbeej, B.; Radom, L.; Smith, D. M. *J. Am. Chem. Soc.* **2012**, *134*, 1591–1599.

(57) Sharma, P. K.; Chu, Z. T.; Olsson, M. H. M.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9661–9666.