THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

Cite This: J. Phys. Chem. A 2019, 123, 5580−5589

pubs.acs.org/JPCA

Are Explicit Solvent Models More Accurate than Implicit Solvent Models? A Case Study on the Menschutkin Reaction

Published as part of The Journal of Physical Chemistry virtual special issue "Leo Radom Festschrift".

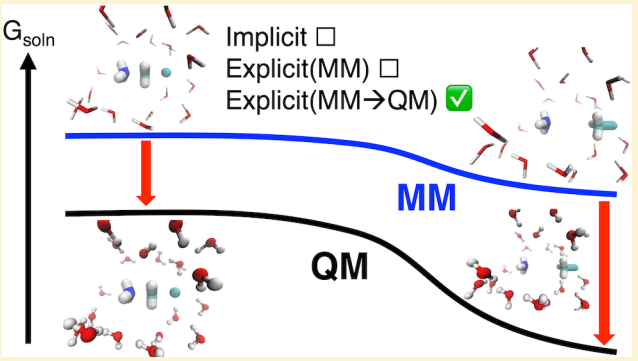
Junbo Chen,[†] Yihan Shao,[‡] and Junming Ho*[,†]

[†]School of Chemistry, University of New South Wales, Sydney, NSW 2052, Australia

[‡]Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019, United States

Supporting Information

ABSTRACT: In this work, contemporary quantum mechanical (QM) implicit solvent models (SMD, SM12, and COSMO-RS) and a molecular mechanical (MM) explicit solvent model were used to predict the aqueous free energy barrier of a simple Menschutkin reaction (NH_3 + CH_3Cl). Surprisingly, the explicit solvent approach performed the worst, while the implicit solvent models yielded reasonably accurate values that are in accord with available experimental data. The origin of the large error in the explicit solvent model was due to the use of a fixed set of Lennard-Jones parameters during the free energy perturbation (FEP) calculations. Further analyses indicate that M06-2X/6-31+G(d,p) yielded solute−solvent interaction energies that are in good agreement with benchmark DLPNO-CCSD(T)/CBS values. When end-state MM to M06-2X/6-31+G(d,p) corrections were added using FEP, it significantly improved the accuracy of the explicit solvent MM result and demonstrated that the accuracy of these models may be systematically improved with end-state corrections based on a validated QM level of theory.



INTRODUCTION

The development of efficient quantum chemistry algorithms and powerful computer architectures has facilitated the use of computation to aid experimental studies. Notably, it is now possible to predict rate and equilibrium constants of gas-phase reactions to accuracies that are comparable to experimental precision. For example, the Pople diagram[1] introduced by John Pople more than 50 years ago allows chemists to make an informed decision about the quality of their calculation and how it can be systematically improved to allow more accurate prediction of reaction outcomes. The Jacob's Ladder provides a similar hierarchy for density functional theory (DFT) methods.[2]

Unfortunately, such a comprehensive framework does not currently exist for solution-phase theoretical chemistry. At present, quantum mechanical (QM) implicit solvent models[3,4] (also known as continuum solvent models) represent the most popular approach for solvation modeling due to their low cost, moderate accuracy, and ease of use. Examples include the polarizable continuum model (PCM) and its variants (e.g., C-PCM[5] and IEF-PCM[6]), the PBSA model in Jaguar,[7,8] COSMO-RS,[9,10] and the Minnesota solvation models (SMD,[11] SM8,[12] and SM12[13]). Many of these models are reliant on the experimental data set used in the parametrization and hence are generally valid only at room temperature for selected neat solvents, and they perform best when applied to molecules that are similar to those in the training set.[14] Typically, these models can predict free energies of solvation to within 5 kJ mol^{-1} for neutral solutes, while the errors for singly charged molecules easily exceed 20 kJ mol^{-1}.[15]

A recent study by one of us has indicated that the free energy barriers for some of the simplest S_N2 reactions estimated by some of these models have errors of at least 20 kJ mol^{-1} compared to experiment.[16] Other groups have also made similar observations for related reactions.[17,18] This is a concern because of the widespread use of these models for elucidating the mechanism of various types of reactions. Unfortunately, the semiempirical nature of implicit solvent models also means that these models are less amenable to systematic improvement, and alternative methods are needed to advance the field. In this regard, explicit solvent models are in principle more general as the description of solute−solvent interactions are modeled explicitly and may be systematically improved using increasingly accurate potentials such as polarizable force fields and hybrid quantum mechanics/molecular mechanics (QM/MM) potentials.

For a small number of model reactions, these methods have been shown to yield kinetic and thermodynamic data that are

Received: April 29, 2019
Revised: June 6, 2019
Published: June 6, 2019

DOI: 10.1021/acs.jpca.9b03995
J. Phys. Chem. A 2019, 123, 5580−5589

in good agreement with experiment. Early work by Jorgensen demonstrated that it is possible to obtain accurate estimates of the activation free energies of S_N2 reactions using the free energy perturbation (FEP) method in conjunction with the OPLS and TIP4P water potentials.[19] This approach has also been applied to various organic reactions,[20−22] including a recent study by Harvey, Sunoj, and co-workers,[23] which used this method to provide an accurate model of the mechanism of the Morita-Baylis-Hillman reaction—a challenging reaction for theoretical investigations due to the large solvation and entropic effects.[18] More recent developments in the field include the implementation of more robust statistical free energy methods (e.g., Bennett's Acceptance Ratio), polarizable and QM/MM potentials,[24−26] and cost-effective reference potential methods[27−31] to accelerate QM/MM calculations.

An interesting observation from these new developments is that the use of supposedly more rigorous potentials does not necessarily translate to gains in accuracy. For example, Mei and co-workers have found that QM/MM potentials did not yield better results compared to MM methods for the calculation of transfer free energies[32] and the free energy profiles of simple organic reactions.[33] Similarly, a recent study has also concluded that QM/MM hydration free energies using a variety of QM methods, such as HF, MP2, and DFT, are generally inferior to purely classical results.[34] This raises the question as to whether the accuracy of explicit solvent models can indeed be systematically improved through the use of more accurate potentials. In particular, these models are at least several orders of magnitude more expensive than implicit solvent models, and it is of interest to examine if the extra cost results in an improvement in accuracy, if any.

In this study, we will address these questions by focusing on a simple Menschutkin reaction (NH_3 + CH_3Cl), which has been studied by many groups, including the Gao,[35] Gordon,[36] Truong,[37] Floris,[38] Chuang,[39] Wiberg,[40] Kato,[41] and Shaik[42] groups, using various electronic structure methods and solvent models. Early work by Gao employed FEP in conjunction with Monte Carlo simulations with MM[35] and AM1/MM[43] potentials to obtain estimates of ∼109 kJ mol^{-1} for the aqueous free energy barrier, while Shaik performed valence bond calculations coupled to a polarizable continuum model to yield a value of 91.2 kJ mol^{-1}.[42] From these studies, the resulting estimates for the barrier ranged from 90.4 to 131.4 kJ mol^{-1}. Since there is no experimental data for this reaction, the experimental value for the related NH_3 + CH_3I reaction (98.3 kJ mol^{-1})[44] is often used to gauge the accuracy of these calculations. However, this value should be regarded as a lower-bound estimate, since the iodide anion is a better leaving group compared to chloride. Calculation of the reaction free energy was not considered because of the large uncertainty in the experimental value (142.3 ± 42 kJ mol^{-1}).[44]

The large scatter in free energy barrier estimates from the above-mentioned theoretical studies is a direct result of the choice of electronic structure methods, MM and QM/MM potentials, and solvent models. Consequently, it is difficult to judge the relative performance of implicit versus explicit solvent models. Since the publication of these studies (the latest one was in 2007 by Shaik), newer-generation solvent models and electronic structure methods have also been developed, and it is timely to revisit this reaction. This paper is set out as follows: We will first determine an accurate level of theory (CCSD(T)/CBS) for the gas-phase barrier. Solvation free energies from QM implicit solvent models (SMD, SM12, and COSMO-RS) and explicit solvent FEP simulations (CHARMM force field (CGENFF) and TIP3P potentials) are used in conjunction with this value to determine the aqueous barrier. Finally, we present a careful study of how (MM→QM) end-state corrections can be applied to improve the accuracy of FEP(MM) calculations.

COMPUTATIONAL DETAILS

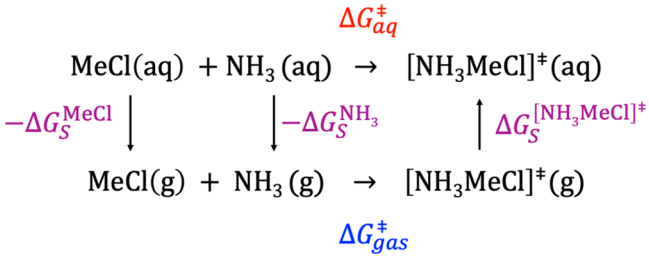
Gas-Phase Calculations. The gas-phase energies were obtained using the standard expression shown in Equation 1. E_{gas} and ZPVE are the gas-phase electronic and zero-point vibrational energies, respectively, and the last two terms represent the thermal correction to the Gibbs free energy.

$$G^{\circ}_{gas} = E_{gas} + \mathrm{ZPVE} + H_{therm} - TS_{tot} \quad (1)$$

In this work, the geometries and thermal corrections were determined at the M06-2X/6-31+G(d,p) level of theory.[45] Single point calculations were performed at the DLPNO-CCSD(T)/CBS,[46,47] B3LYP,[48,49] ωB97M-V,[50] and ωB97X-D/6-31+G(d,p) levels of theory.[51] For the complete basis set (CBS) extrapolation of the CCSD(T) electronic energies, we used aug-cc-pVDZ and aug-cc-pVTZ basis sets and the optimum values for α and β determined by Neese and co-workers.[52] The superscript "o" denotes that these quantities are calculated at a reference state of 1 atm.

Implicit Solvent Calculations. The above gas-phase energies were combined with solvation free energies from the SMD,[11] SM12,[13] and COSMO-RS[9,10] implicit solvent models (according to the thermodynamic cycle in Scheme 1) to yield the aqueous Gibbs free energy barrier (ΔG^{*}_{aq}).

Scheme 1. Thermodynamic Cycle for the Calculation of Aqueous Barrier



$$\Delta G^{*}_{aq} = \Delta G^{\circ}_{gas} + \Delta G^{*}_{S}(\mathrm{TS}) - \Delta G^{*}_{S}(\mathrm{MeCl}) - \Delta G^{*}_{S}(\mathrm{NH_3}) - \Delta G^{\circ\rightarrow *} \quad (2)$$

The superscript "*" denotes that the properties were calculated with respect to a reference state of 1 mol L^{-1}, and a correction ($\Delta G^{\circ\rightarrow *}$ = 7.92 kJ mol^{-1}) was added to account for the difference in the gas-phase and solution-phase standard states. Normally for small rigid solutes, the solvation free energies are calculated via eq 3, where the solution-phase (E_s) and gas-phase (E_g) energies are evaluated on the same geometry optimized in the gas or solution phase ($\mathbf{R}_g$ and $\mathbf{R}_s$).

$$\Delta G^{*}_{S}(\mathbf{R}_s) = E_s(\mathbf{R}_s) + G_{nes}(\mathbf{R}_s) - E_g(\mathbf{R}_s) \quad (3)$$

In cases where the solution-phase geometry ($\mathbf{R}_s$) differs appreciably from that of the gas phase ($\mathbf{R}_g$), the use of Equation 4 is recommended, because it accounts for the effect of geometrical relaxation in the solvation free energy. G_{nes}

DOI: 10.1021/acs.jpca.9b03995
J. Phys. Chem. A 2019, 123, 5580−5589

refers to the nonelectrostatic contributions to the free energy of solvation. As we will see later, the effect of geometrical relaxation is rather large for the Menschutkin transition state, and all implicit solvent calculations in this work are based on Equation 4.

$$\Delta G_S^*(\mathbf{R}_s, \mathbf{R}_g) = E_s(\mathbf{R}_s) + G_{nes}(\mathbf{R}_s) - E_g(\mathbf{R}_g) \quad (4)$$

The aqueous-phase geometries were obtained at the SMD-M06-2X/6-31+G(d,p) level. To adhere to their parametrization protocol, the SMD and SM12 models were applied at the M06-2X/6-31+G(d,p) level of theory, while the COSMO-RS model was applied at the BP/TZP level.

Explicit Solvent Calculations. The reaction change in solvation free energy $\Delta\Delta G_S^*$ was also evaluated using an FEP approach developed by Jorgensen and co-workers, which is described in detail elsewhere.[24,25,53] As noted in Equation 2, $\Delta\Delta G_S^*$ is simply the difference in free energies of solvation of the transition state (TS) and that of the reactants: ΔG_S^* (TS) − ΔG_S^* (MeCl) − ΔG_S^* (NH_3). Briefly, this approach involved a selection of a starting (X) and final (Y) structure and a large number of intermediate structures obtained by linear interpolation. Alternatively, a minimum energy reaction path (MERP) can also be traced from the transition state to the reactants. The key point here is that the geometric change of a distance (or angle) must be small to ensure fast convergence of the perturbation. As shown in Figure 1, $E_{int}(\mathbf{x}_i)$ refers to the

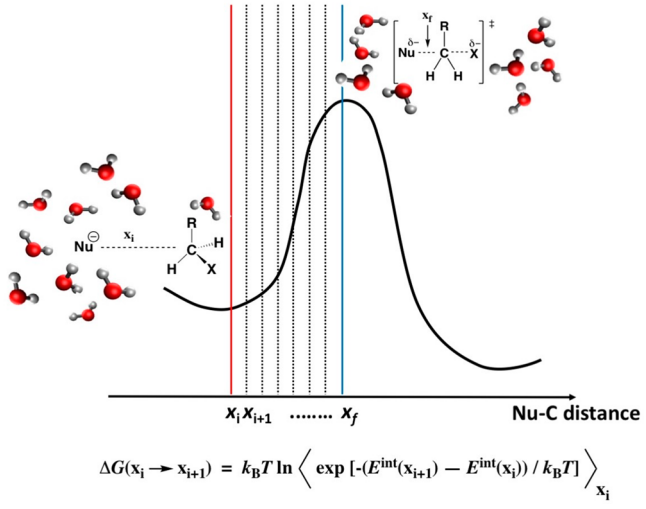


Figure 1. Illustration of Jorgensen's FEP method for calculating $\Delta\Delta G_S$ along a reaction coordinate. Only the nearest solvent molecules are shown.

interaction energy between the solute (in geometry $\mathbf{x}_i$) and the solvent molecules, and the free energy change in solvation along each step $\Delta G(x_i \rightarrow x_{i+1})$ can be obtained via the Zwanzig equation.[54] The $\Delta\Delta G_S^*$ can then be obtained from the sum of these individual free energy changes.

Following the procedure in a recent study by Harvey and Sunoj,[23] we started with the SMD transition-state geometry and a reactant configuration, where the distance between the ammonia (N atom) and methyl chloride (C atom) is 10 Å. A total of 100 intermediate geometries were generated by linear interpolation. For each structure, the solute was placed into a rectangular box (x = 30 Å, y = 30 Å, and z = 40 Å) filled with 1200 TIP3P water molecules corresponding to the density of water 1 g cm^{-3}. The general CHARMM atom types[55,56] were assigned to methyl chloride (using the cgenff.paramchem.org server) and the solute geometry was held fixed *held fixed* during the molecular dynamics (MD) simulation, so that no intramolecular MM terms were used. The atomic charges for each structure were replaced by the Merz−Kollman electrostatic potential (ESP) charges obtained using the SMD model and M06-2X/6-31+G(d,p) level of theory; the methyl chloride van der Waals parameters were kept fixed for all structures. A 5 ns production run was launched after 1 ns of equilibration under NPT ensemble (25 °C and 1 atm) and periodic boundary conditions. A total of 5000 snapshots were sampled from each MD trajectory, and the interaction energy between the solute and solvent was evaluated under periodic boundary conditions. Both forward and reverse FEP simulations were performed to assess convergence, and the difference in $\Delta\Delta G_S^*$ was less than 2 kJ mol^{-1} (see Figure S1).

This $\Delta\Delta G_S^*$ is then combined with gas-phase barrier (ΔG^o_{gas}) to estimate the solution-phase free energy (Scheme 1). A subtle point here is that, because the FEP $\Delta\Delta G_S$ is determined using SMD geometries for the transition state and reactant, it does not include the effect of geometrical relaxation (cf. eq 3). As such, the corresponding gas-phase barrier is calculated using CCSD(T)/CBS energies on SMD geometries (ΔG^o_{gas}//SMD) and thermal corrections to derive the aqueous-phase barrier. This barrier (157.3 kJ mol^{-1}) is significantly lower than the calculated value on gas-phase geometries (201.2 kJ mol^{-1}), because the aqueous-phase transition state is early relative to that of the gas phase (vide infra).

All electronic structure calculations were performed using the Gaussian16,[57] Q-CHEM,[58] ADF,[59−61] and ORCA[62,63] programs. The COSMO-RS[64] calculations were performed using ADF, and the SM12 (Merz−Kollman charges) calculations were performed in Q-CHEM. The DLPNO−CCSD(T) calculations were performed using ORCA, and other electronic structure calculations were performed in Gaussian16. The MD simulations were performed using NAMD,[65] and VMD[66] was used for visualization and analysis of the trajectories.

RESULTS AND DISCUSSION

Gas-Phase Barrier. Figure 2 compares the gas-phase barrier (ΔG^o_{gas}) obtained at different levels of theory. As shown, the barriers obtained using the small basis set DFT calculations differ appreciably from the CCSD(T)/CBS values (201.2 kJ mol^{-1}). Notably, the B3LYP/6-31+G(d,p) value differs by more than 20 kJ mol^{-1}, and increasing the basis set to 6-311+G(3df,2p) resulted in a modest improvement in ΔG^o_{gas} (by less than 4 kJ mol^{-1}). This is interesting, since many of the previous studies on the Menschutkin reaction are based on a small basis set DFT or MP4 calculations. For example, Truong et al. calculated the gas-phase free energy barrier to be 189 and 191 kJ mol^{-1} for the MP4(SDTQ) and BH&HLYP/6-31G(d,p) levels of theory,[37] while Castejon and Wiberg obtained a value of 175 kJ mol^{-1} at the B3LYP/6-31+G(d) level.[40] These deviations raise the question as to whether the good agreement with experiment reported by earlier studies was due to systematic cancellation of errors from the gas-phase and solvation calculations.

Solution-Phase Barrier. Using the CCSD(T)/CBS gas-phase barrier of 201.2 kJ mol^{-1}, we determined the aqueous-phase barrier using the SMD, SM12, and COSMO-RS models (Figure 3). As shown, the implicit solvent estimates of the

DOI: 10.1021/acs.jpca.9b03995
J. Phys. Chem. A 2019, 123, 5580−5589

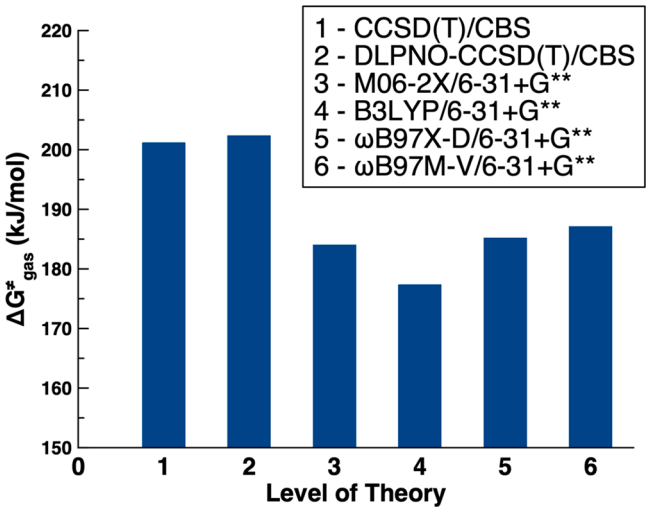
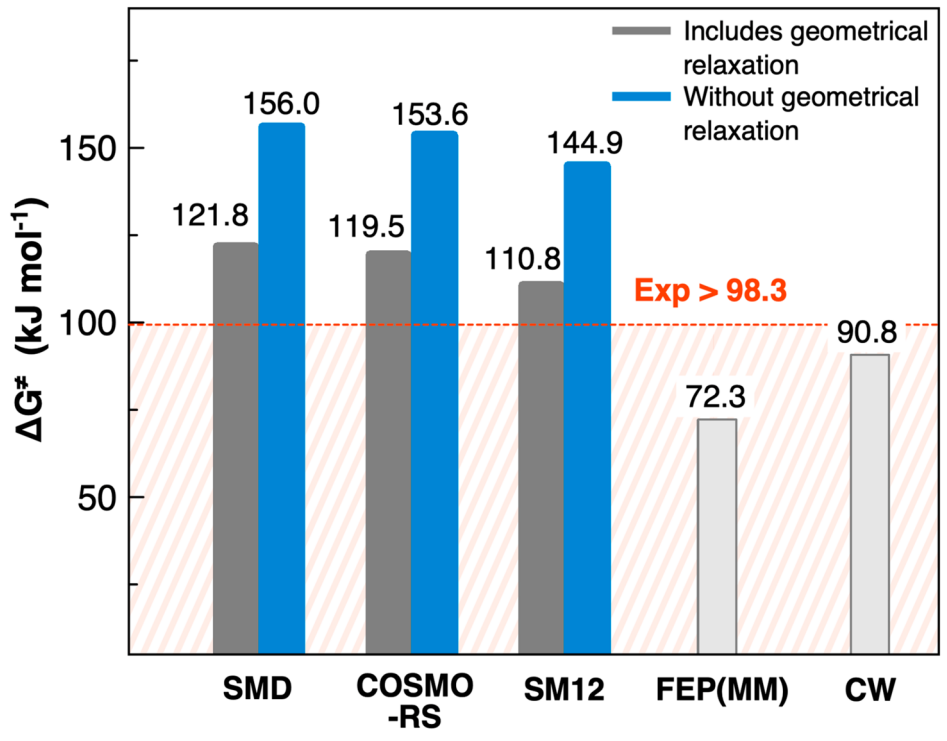


Figure 2. Gas-phase activation barrier for the Menshutkin reaction computed at different levels of theory.

barriers are dramatically improved (gray bars) when the solvation free energies were calculated using Equation 4 compared to values where the gas and solution-phase energies were determined on the same geometry (Equation 3). This is because the solution-phase transition state is "early" compared to the gas-phase geometry (Figure 3b). In particular, the C−N bond distance is ∼0.3 Å longer compared to corresponding gas-phase structure, which is also consistent with results from previous theoretical studies.[36−38,40] As such, it is important to account for this geometrical contribution in the solvation free energy calculation, which amounts to ∼34 kJ mol^{-1}. This leads to aqueous barriers that range from 110.8 to 121.8 kJ mol^{-1}, which are in the right ballpark compared to the experiment estimate of more than 98.3 kJ mol^{-1}.

For the explicit solvent calculation, a total of three independent FEP runs were performed, and for each run, the resulting $\Delta\Delta G_S^*$ value was obtained from the average of the forward and reverse runs to yield values of −76.3, −78.3, and −76.7 kJ mol^{-1}. Accordingly, the FEP(MM) $\Delta\Delta G_S^*$ value is estimated to be −77.1 ± 2.2 kJ mol^{-1} (error bar corresponds to 2σ). As explained in the Methods section, because the FEP(MM) $\Delta\Delta G_S^*$ value was obtained from SMD geometries for the transition state and reactant, it does not include the effect of geometrical relaxation. As such, the corresponding gas-phase barrier is calculated using CCSD(T)/CBS energies on SMD geometries and thermal corrections (ΔG°_{gas}//SMD) to derive the aqueous-phase barrier. This value is ∼46 kJ mol^{-1} lower compared to the values computed on gas-phase geometries (Figure 2), and when combined with the

(a)



(b)

	r(C-N)/Å	r(C-Cl)/Å	H-C-Cl angle/°
M06-2X/6-31+G(d,p) (Rg)	1.813	2.448	81.6
SMD @ M06-2X/6-31+G(d,p) (Rs)	2.122	2.206	92.9

Figure 3. (a) The solution-phase reaction barrier (ΔG_{aq}^*) for the Menshutkin reaction calculated using different solvent models. CW refers to the FEP value reported by Castejon and Wiberg (ref 40). The experimental value for the aqueous activation barrier for NH_3 + CH_3I is shown. (b) Selected geometrical parameters for the gas-phase and SMD transition-state structures.

DOI: 10.1021/acs.jpca.9b03995
J. Phys. Chem. A 2019, 123, 5580−5589

FEP(MM) $\Delta\Delta G^{\ddagger}_{s}$ we obtained a value of 72.3 kJ mol^{-1} for the aqueous barrier. This deviates by more than 25 kJ mol^{-1} from the experimental estimate, which is surprising given the good performance displayed by the implicit solvent models.

Castejon and Wiberg had employed a similar FEP procedure and estimated the aqueous free energy barrier to be 90.8 kJ mol^{-1}.[40] The difference between their value and our calculation may be attributed to their use of the gas-phase reaction coordinate, which is a source of error, because the aqueous-phase transition-state geometry is significantly early relative to the gas-phase geometry (Figure 3b). Additionally, Castejon and Wiberg calculated the gas-phase barrier using B3LYP, which is ∼25 kJ mol^{-1} lower than the CCSD(T)/CBS benchmark. As such, it is likely that their calculation agreed better with experiment due to fortuitous cancellation of errors. Thus, an advantage of computing the solution-phase free energy change from separate gas-phase and solvation contributions (eq 2) is that one can use very accurate electronic structure methods to estimate the gas-phase energies thereby providing a clearer perspective of the errors underlying the solvent models and how to improve them.

MM Versus QM Interaction Energies. To probe the origin of the errors in the FEP(MM) barrier, we first performed a benchmarking study to determine an accurate level of theory for modeling the solute−solvent interaction energies. For this purpose, we sampled the MD trajectories of the reactants and transition state to obtain 50 frames of the solute solvated by the nearest 10 TIP3P water molecules in various configurations. For each snapshot, we further determined the interaction energy between the solute and solvent at the DLPNO−CCSD(T)/CBS level and used this to benchmark several DFT methods. Table 1 summarizes the performance of these methods.

Table 1. Mean and Maximum Deviation (in kJ mol^{-1}) in DFT/6-31+G(d,p) Interaction Energies Compared to DLPNO−CCSD(T)/CBS Values for 10-Water Clusters

TS·$(H_2O)_{10}$	B3LYP-D3[a]	ωB97X-D[a]	M06-2X[a]	M06-2X-D3[a]
MSD[b]	−3.6	5.0	0.0	−3.9
MAD	4.4	5.3	3.4	4.6
ADmax	12.6	13.5	9.5	13.4
R·$(H_2O)_{10}$				
MSD[b]	−5.2	2.6	−0.9	−2.9
MAD	5.3	3.6	3.3	3.9
ADmax	10.3	12.3	9.8	11.8

[a]Includes counterpoise correction.[67,68] [b]Mean signed deviation (DFT minus DLPNO−CCSD(T)/CBS interaction energies).

As shown, the mean signed deviation (MSD) indicates that the B3LYP-D3, M06-2X-D3, and ωB97X-D methods systematically under- or overestimate the interaction energy, while M06-2X performed best with mean absolute deviation (MAD) of less than 4 kJ mol^{-1} for both the reactant and transition-state water clusters. As such, this level of theory is used to determine the interaction energies of *larger* water clusters containing the solute and 40 nearest TIP3P water molecules to assess the accuracy of the MM force field. The XY scatter plots for the MM and M06-2X interaction energies are presented in Figure 4. Note that in this case the DFT interaction energies are not counterpoise-corrected, because MM (CGENFF/TIP3P) interaction energies do not account for basis set superposition errors.[67,68] According to these results, the origin

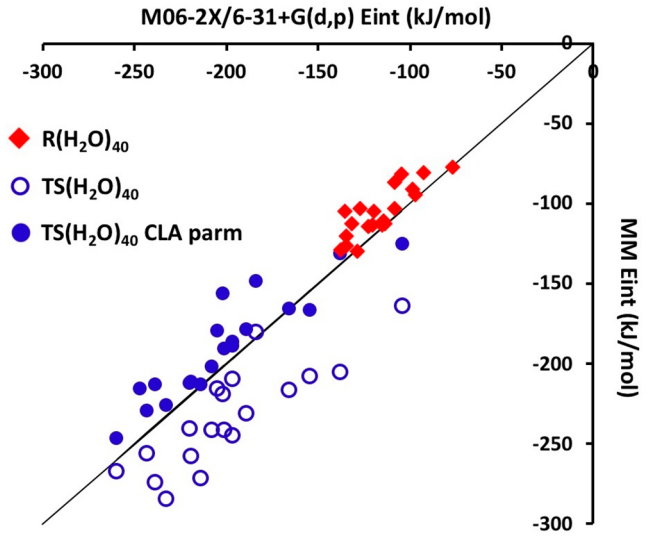


Figure 4. Scatter plot for the MM vs M06-2X/6-31+G(d,p) interaction energies between the solute and 40 nearest TIP3P water molecules. CLA parm refers to MM interaction energies computed using chloride vDW parameters for the chlorine atom in the transition state.

of the errors in the FEP(MM) solution-phase barrier appears to arise as a result of overestimation of the solvent transition-state interaction energy (more negative) compared to the reactants. In particular, the mean absolute deviation between the MM and DFT interaction energies is ∼36 and 11 kJ mol^{-1} for the transition state versus reactant, respectively.

One possible explanation for the larger errors observed at the transition state is that (unlike the atomic charges) the methyl chloride van der Waals (vDW) parameters are kept fixed for all the interpolated geometries in the FEP simulations. For this reason, the vDW parameter for the chlorine atom in the transition state is identical to that in the reactant (methyl chloride; CGENFF atom type CLGA1). This is a potential source of error, because there is significant charge development at the transition state (ca. 0.5e increase in ESP atomic charge at the chlorine atom relative to methyl chloride), such that the chlorine resembles a chloride anion in the transition state. To determine if this was indeed the origin of the large CGENFF/TIP3P errors for the transition state, the MM interaction energies were recomputed using vDW parameters for the chloride anion (CHARMM atom type CLA), while the remaining atoms are unchanged. Interestingly, this reduces the mean absolute error (relative to M06-2X/6-31+G(d,p) values) from 36 to 15 kJ mol^{-1}, which is now much closer to the CGENFF/TIP3P error for the reactant (11 kJ mol^{-1}).

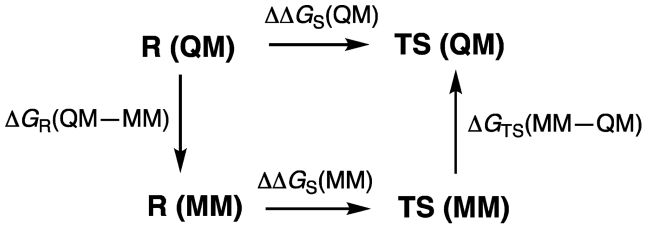
This highlights a potential problem with the current FEP procedure of using fixed vDW parameters along the reaction coordinate, as the vDW contributions can be rather large (−20 kJ mol^{-1}) in this case. One possible remedy for this problem is to use chloride (CLA) and chlorine atom (CLGA1) vDW parameters for the transition state and reactant, respectively, and the vDW parameter at the intermediate geometries may be obtained from interpolation between the two end points. Another approach is to employ MM → QM corrections at the end points,[27] and we examine the use of this strategy in the next section.

Improving the Accuracy of FEP Results via End-State Corrections. Performing the MD simulations at the MM

DOI: 10.1021/acs.jpca.9b03995
J. Phys. Chem. A 2019, 123, 5580−5589

level, followed by FEP from MM to QM (or QM/MM) between various states was first suggested by Warshel et al.,[29,30,69] and it has been used by several other groups.[27,28,70−74] For example, a recent study by Mei and co-workers employed free energy perturbation to obtain MM to QM/MM corrections at the two end states to improve transfer free energy calculations of organic solutes in water and chloroform.[32] The use of interaction energy difference E_{int}(QM) − E_{int}(MM) has been shown to lead to well-converged results, since the internal degrees of freedom are not considered.[75−78] The fixed solute geometry in our FEP method should further accelerate convergence. In this work, we employ a similar approach (Scheme 2) to improve the accuracy of $\Delta\Delta G_S^*$(MM) by performing FEP calculations of the interaction energy difference between CGENFF/TIP3P and M06-2X/6-31+G(d,p).

Scheme 2. Thermodynamic Cycle for the Calculation of $\Delta\Delta G_S$(QM) through End-State Corrections



$$\Delta\Delta G_S(QM) = \Delta\Delta G_S(MM) + \Delta G_{TS}(MM\text{-}QM) + \Delta G_R(QM{-}MM)$$

The deviations in QM and MM solute−solvent interaction energies are expected to arise primarily from the description of the solute interaction with the first few solvent shells, and it is therefore unnecessary to consider the entire solvent box for the calculation of QM interactions. Accordingly, we performed a systematic study to determine the optimal solvated cluster for calculating the end-state corrections. In particular, we randomly selected 20 frames from the MD trajectory of the transition state and that of the reactant and examined how the interaction energy difference ΔE_{int} = E_{int}(QM) − E_{int}(MM) varies as a function of the number of nearest solvent molecules. Table 2 provides the mean absolute values of ΔE_{int} as the number of nearest water molecules increases from 10 to 160. As shown, the difference in interaction energy ceases to change by more than a few kilojoules per mole, once the cluster size reaches 40 water molecules. As such, the FEP MM→QM corrections were determined using this cluster size for the reactants and products.

Table 2. Mean Absolute Deviation (in kJ mol^{-1}) in CGENFF/TIP3P and M06-2X/6-31+G(d,p) Interaction Energies as a Function of the Number of Water Molecules in the Cluster[a]

$n(H_2O)$	R	TS
10	6.5	50.6
20	7.4	40.1
40	10.7	35.6
80	10.5	38.3
160	7.1	40.4

[a]Averaged over 20 frames, and M06-2X/6-31+G(d,p) values do not include counterpoise correction (see text).

As noted above, these M06-2X/6-31+G(d,p) interaction energies do not include counterpoise corrections. For the 40-water clusters, the counterpoise corrections for the reactant and transition states are very similar, amounting to 18 and 17 kJ mol^{-1} on average. On the basis of the thermodynamic cycle in Scheme 2, we expect these corrections would mostly cancel in the calculation of $\Delta\Delta G_S^*$(QM). As such, all DFT interaction energies do not include corrections for basis set superposition error.

Shown in Figure 5 are the end-state corrections (MM to QM) as a function of the number of frames included in the exponential average. As shown, these corrections appear to converge after ∼2000 frames, where the subsequent values are all within a few kilojoules per mole. In particular, the MM→QM corrections based on the average of the last three data points (2000, 2500, and 3000 frames) are 2.9 and −34.4 kJ mol^{-1} for the transition state and reactants, respectively. Interestingly, these corrections are in contrast to what one would expect from Figure 4. Notably, the error in the MM interaction energy is considerably higher for the transition state (36 kJ mol^{-1}) as compared to the reactant (11 kJ mol^{-1}), which is the opposite of the calculated MM→QM corrections. This is presumably because the configurations that give rise to these large errors contribute very little to the exponential average, that is, are statistically less important on the QM surface. These values combined with the original $\Delta\Delta G_S^*$(MM) of −77.1 kJ mol^{-1} amounts to a $\Delta\Delta G_S^*$(QM) of −39.8 kJ mol^{-1}. When this value is combined with the gas-phase barrier ΔG°_{gas} of 157.3 kJ mol^{-1}, the calculated aqueous-phase barrier is now 109.6 kJ mol^{-1}, which is in closer agreement to the estimates from the implicit solvent models.

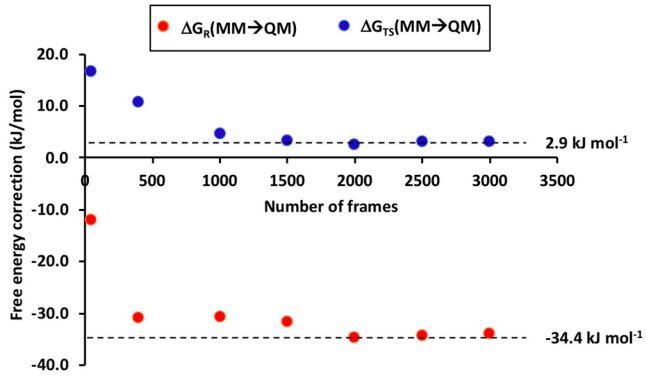


Figure 5. End-state (MM→QM) corrections for reactant and transition state as a function of the number of sampled frames.

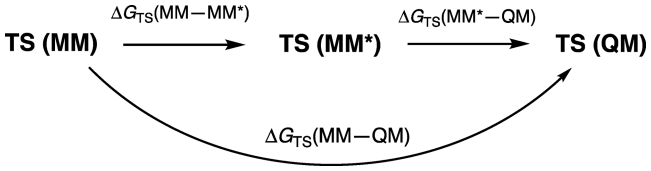
Because the FEP corrections were computed for frames sampled during the MM simulation it is considered a one-sided method as compared to two-sided methods (forward and reverse) such as simple-overlap sampling (SOS) or Bennett's Acceptance Ratio (BAR) that are considered more robust.[79] Unfortunately, two-sided methods entail MD simulations at the DFT level, which are computationally very expensive. While hybrid quantum mechanics/molecular mechanics (QM/MM) potentials are frequently used in place of ab initio MD simulations (AIMD), there are also complications associated with these methods, such as the optimal QM region size (number of solvent molecules to include in the QM region)

DOI: 10.1021/acs.jpca.9b03995
J. Phys. Chem. A 2019, 123, 5580−5589

and choice of MM force field, which can significantly influence the accuracy of the QM/MM simulations.[80]

Alternatively, several measures such as the reweighting entropy,[32] Kish's effective sampling size,[81] and analysis of the standard deviation[82,83] in the distribution of E_{int}(QM) − E_{int}(MM) have been proposed to assess the quality of FEP calculations. We calculated the reweighting entropy (S) associated with the end-state FEP(MM→QM) corrections, which are 0.32 and 0.2 for the transition state and reactant, respectively. S has a maximum value of 1.0, where all frames contribute equally to the exponential average, so the FEP correction generally becomes more reliable as the reweighting entropy increases. Because our S values are relatively low compared to the recommended value of 0.65,[32] we adopted another approach to check the robustness of our MM→QM corrections. Specifically, we considered an alternative pathway as shown in Scheme 3, where the correction for the transition

Scheme 3. An Alternative Two-Step Pathway to Obtain the ΔG_{TS}(MM→QM) Correction



state is obtained from two steps, that is, ΔG_{TS}(MM→QM) = ΔG_{TS}(MM→MM*) + ΔG_{TS}(MM*→QM). In Scheme 3, MM* denotes that the vDW parameters for the CGENFF chloride atom (CLGA1) are replaced by the CHARMM parameters for chloride (CLA). For the ΔG_{TS}(MM→MM*) correction, we performed both the forward and reverse FEP calculations and obtained values of 25.3 and −24.2 kJ mol^{-1}, respectively. The good agreement between these values indicates that the FEP ΔG_{TS}(MM→MM*) calculation is well-converged. The ΔG_{TS}(MM*→QM) is calculated in the same manner as described above but using configurations sampled from the trajectory employing MM* parameters, and we obtained a value of −26.6 kJ mol^{-1}. Accordingly, the two-step estimate of ΔG_{TS}(MM→QM) amounts to −1.8 kJ mol^{-1}, which agrees well with the one-step estimate of 2.9 kJ mol^{-1}. These results indicate that our FEP end-state corrections are reasonably robust, and the calculated aqueous barrier of 109.6 kJ mol^{-1} has an error bar of ca. ±5 kJ mol^{-1}.

CONCLUSIONS

In this paper, we examined the performance of various QM implicit solvent models and an explicit solvent model for the calculation of the aqueous barrier of a simple Menschutkin reaction. To our surprise, explicit solvent FEP simulations with CGENFF/TIP3P potential significantly underestimated the barrier by at least 25 kJ mol^{-1}, while the implicit solvent model values are in reasonable agreement with experiment. Our investigations indicate that the origin of this large deviation may be a result of using a fixed set of vDW parameters during the FEP simulations along the reaction path.

The results from this study indicate that (MM→QM) end-state corrections provide an effective strategy to improve the accuracy of the FEP(MM) simulations of $\Delta\Delta G_S^*$. Notably, by considering errors in the MM interaction energies between the solute and nearest 40 water molecules, we were able to increase the $\Delta\Delta G_S^*$ by ∼38 kJ mol^{-1}, bringing the FEP barrier to ∼110 ± 5 kJ mol^{-1}. As noted in our discussion, the actual error is likely to be smaller, as the experimental barrier is likely to be significantly higher than 98 kJ mol^{-1}. This is because the value of 98 kJ mol^{-1} is based on the experimental barrier for ammonia with methyl iodide, which is expected to be lower compared to the reaction with methyl chloride.

There are several advantages in the thermodynamic cycle-FEP procedure presented in this work. The partitioning of the solution-phase free energy change into gas-phase and solvation contributions (Scheme 1) allows one to use accurate quantum-chemistry calculations for the gas-phase energies and systematically improve the accuracy of the solvation contribution. The FEP(MM) simulations of $\Delta\Delta G_S^*$ employing fixed solute geometry is easy to implement, and the MM→QM end-state corrections are significantly cheaper compared to direct QM/MM molecular dynamics simulations. This strategy is expected to work well most of the time, except in difficult cases, where the transition-state structure from an implicit solvent model might deviate substantially from the one from explicit solvent calculations. Finally, note that the current (equilibrium) approach invokes an important assumption that solvent relaxation is fast relative to reaction coordinate motion. However, for chemical reactions with short-lived transition states this adiabatic assumption may not be correct. Transition path sampling (TPS) methods[84−86] would account for these effects, and it would be interesting to examine how much TPS results would differ from equilibrium free energy pathway simulations.

Future work will examine the use of more systematic approaches such as force matching[87] to enhance the overlap between the MM and QM phase space, and also more efficient methods such as linear-scaling fragmentation methods to further speed up the calculation of MM→QM end-state corrections. Work is also underway toward a more extensive study of different reaction types and solvents to ascertain the broader accuracy of our proposed procedure.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.9b03995.

Tables of calculated gas-phase and solution-phase energies, XY scatterplots of MM versus QM interaction energies, and Cartesian coordinates of M06-2X optimized geometries (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: junming.ho@unsw.edu.au.

ORCID

Junbo Chen: 0000-0002-0714-6710
Yihan Shao: 0000-0001-9337-341X
Junming Ho: 0000-0001-9381-924X

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

J.H. thanks the Australian Research Council for funding (DE160100807) and the Australian National Computational Infrastructure, UNSW, Intersect NSW, and Pawsey Super-

DOI: 10.1021/acs.jpca.9b03995
J. Phys. Chem. A 2019, 123, 5580−5589

computing Centre for generous allocation of computing resources.

REFERENCES

(1) Pople, J. A. Two-Dimensional Chart of Quantum Chemistry. *J. Chem. Phys.* **1965**, *43*, S229−S230.
(2) Perdew, J. P.; Schmidt, K. Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy. *AIP Conf. Proc.* **2000**, *577*, 1−20.
(3) Cramer, C. J.; Truhlar, D. G. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* **1999**, *99*, 2161−2200.
(4) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999−3094.
(5) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, Structures, and Electronic Properties of Molecules in Solution with the C-PCM Solvation Model. *J. Comput. Chem.* **2003**, *24*, 669−681.
(6) Tomasi, J.; Mennucci, B.; Cancès, E. The IEF Version of the PCM Solvation Method: An Overview of a New Method Addressed to Study Molecular Solutes at the QM Ab Initio Level. *J. Mol. Struct.: THEOCHEM* **1999**, *464*, 211−226.
(7) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Honig, B.; Ringnalda, M.; Goddard, W. A. Accurate First Principles Calculation of Molecular Charge Distributions and Solvation Energies from Ab Initio Quantum Mechanics and Continuum Dielectric Theory. *J. Am. Chem. Soc.* **1994**, *116*, 11875−11882.
(8) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. New Model for Calculation of Solvation Free Energies: Correction of Self-Consistent Reaction Field Continuum Dielectric Theory for Short-Range Hydrogen-Bonding Effects. *J. Phys. Chem.* **1996**, *100*, 11775−11788.
(9) Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. Refinement and Parametrization of COSMO-RS. *J. Phys. Chem. A* **1998**, *102*, 5074−5085.
(10) Klamt, A. The COSMO and COSMO-RS Solvation Models. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, No. e1338.
(11) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378−6396.
(12) Cramer, C. J.; Truhlar, D. G.; Marenich, A. V.; Kelly, C. P.; Olson, R. M. Self-Consistent Reaction Field Model for Aqueous and Nonaqueous Solutions Based on Accurate Polarized Partial Charges. *J. Chem. Theory Comput.* **2007**, *3*, 2011−2033.
(13) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Generalized Born Solvation Model SM12. *J. Chem. Theory Comput.* **2013**, *9*, 609−620.
(14) Ho, J.; Coote, M. L. A Universal Approach for Continuum Solvent PKa Calculations: Are We There Yet? *Theor. Chem. Acc.* **2010**, *125*, 3−21.
(15) Cramer, C. J.; Truhlar, D. G. A Universal Approach to Solvation Modeling. *Acc. Chem. Res.* **2008**, *41*, 760−768.
(16) Ho, J.; Ertem, M. Z. Calculating Free Energy Changes in Continuum Solvation Models. *J. Phys. Chem. B* **2016**, *120*, 1319−1329.
(17) Miguel, E. L. M.; Santos, C. I. L.; Silva, C. M.; Pliego, J. R., Jr How Accurate Is the SMD Model for Predicting Free Energy Barriers for Nucleophilic Substitution Reactions in Polar Protic and Dipolar Aprotic Solvents? *J. Braz. Chem. Soc.* **2016**, *27*, 2055−2061.
(18) Plata, R. E.; Singleton, D. A. A Case Study of the Mechanism of Alcohol-Mediated Morita Baylis-Hillman Reactions. the Importance of Experimental Observations. *J. Am. Chem. Soc.* **2015**, *137*, 3811−3826.
(19) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. Theoretical Examination of the SN2 Reaction Involving Chloride Ion and Methyl Chloride in the Gas Phase and Aqueous Solution. *J. Am. Chem. Soc.* **1985**, *107*, 154−163.
(20) Vayner, G.; Houk, K. N.; Jorgensen, W. L.; Brauman, J. I. Steric Retardation of SN2 Reactions in the Gas Phase and Solution. *J. Am. Chem. Soc.* **2004**, *126*, 9054−9058.
(21) Gunaydin, H.; Acevedo, O.; Jorgensen, W. L.; Houk, K. N. Computation of Accurate Activation Barriers for Methyl-Transfer Reactions of Sulfonium and Ammonium Salts in Aqueous Solution. *J. Chem. Theory Comput.* **2007**, *3*, 1028−1035.
(22) Chen, X.; Regan, C. K.; Craig, S. L.; Krenske, E. H.; Houk, K. N.; Jorgensen, W. L.; Brauman, J. I. Steric and Solvation Effects in Ionic SN2 Reactions. *J. Am. Chem. Soc.* **2009**, *131*, 16162−16170.
(23) Liu, Z.; Patel, C.; Harvey, J. N.; Sunoj, R. B. Mechanism and Reactivity in the Morita-Baylis-Hillman Reaction: The Challenge of Accurate Computations. *Phys. Chem. Chem. Phys.* **2017**, *19*, 30647−30657.
(24) Acevedo, O.; Jorgensen, W. L. Quantum and Molecular Mechanical Monte Carlo Techniques for Modeling Condensed-Phase Reactions. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 422−435.
(25) Acevedo, O.; Jorgensen, W. L. Exploring Solvent Effects upon the Menshutkin Reaction Using a Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 8425−8430.
(26) Acevedo, O.; Jorgensen, W. L. Advances in Quantum and Molecular Mechanical (QM/MM) Simulations for Organic and Enzymatic Reactions. *Acc. Chem. Res.* **2010**, *43*, 142−151.
(27) Rod, T. H.; Ryde, U. Quantum Mechanical Free Energy Barrier for an Enzymatic Reaction. *Phys. Rev. Lett.* **2005**, *94*, 138302.
(28) Rod, T. H.; Ryde, U. Accurate QM/MM Free Energy Calculations of Enzyme Reactions: Methylation by Catechol O-Methyltransferase. *J. Chem. Theory Comput.* **2005**, *1*, 1240−1251.
(29) Muller, R. P.; Warshel, A. Ab Initio Calculations of Free Energy Barriers for Chemical Reactions in Solution. *J. Phys. Chem.* **1995**, *99*, 17516−17524.
(30) Štrajbl, M.; Hong, G.; Warshel, A. Ab Initio QM/MM Simulation with Proper Sampling: "First Principle" Calculations of the Free Energy of the Autodissociation of Water in Aqueous Solution. *J. Phys. Chem. B* **2002**, *106*, 13333−13343.
(31) Duarte, F.; Amrein, B. A.; Blaha-Nelson, D.; Kamerlin, S. C. L. Recent Advances in QM/MM Free Energy Calculations Using Reference Potentials. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850*, 954−965.
(32) Wang, M.; Li, P.; Jia, X.; Liu, W.; Shao, Y.; Hu, W.; Zheng, J.; Brooks, B. R.; Mei, Y. Efficient Strategy for the Calculation of Solvation Free Energies in Water and Chloroform at the Quantum Mechanical/Molecular Mechanical Level. *J. Chem. Inf. Model.* **2017**, *57*, 2476−2489.
(33) Li, P.; Jia, X.; Pan, X.; Shao, Y.; Mei, Y. Accelerated Computation of Free Energy Profile at Ab Initio Quantum Mechanical/Molecular Mechanics Accuracy via a Semi-Empirical Reference Potential. I. Weighted Thermodynamics Perturbation. *J. Chem. Theory Comput.* **2018**, *14*, 5583.
(34) König, G.; Pickard, F.; Huang, J.; Thiel, W.; MacKerell, A.; Brooks, B.; York, D. A Comparison of QM/MM Simulations with and without the Drude Oscillator Model Based on Hydration Free Energies of Simple Solutes. *Molecules* **2018**, *23*, 2695.
(35) Gao, J. A Priori Computation of a Solvent-Enhanced SN2 Reaction Profile in Water: The Menshutkin Reaction. *J. Am. Chem. Soc.* **1991**, *113*, 7796−7797.
(36) Webb, S. P.; Gordon, M. S. Solvation of the Menshutkin Reaction: A Rigorous Test of the Effective Fragment Method. *J. Phys. Chem. A* **1999**, *103*, 1265−1273.
(37) Truong, T. N.; Truong, T. T. T.; Stefanovich, E. V. A General Methodology for Quantum Modeling of Free-Energy Profile of Reactions in Solution: An Application to the Menshutkin NH3+CH3Cl Reaction in Water. *J. Chem. Phys.* **1997**, *107*, 1881−1889.
(38) Amovilli, C.; Mennucci, B.; Floris, F. M. MCSCF Study of the S_N2 Menshutkin Reaction in Aqueous Solution within the Polarizable Continuum Model. *J. Phys. Chem. B* **1998**, *102*, 3023−3028.

DOI: 10.1021/acs.jpca.9b03995
J. Phys. Chem. A 2019, 123, 5580−5589

(39) Chuang, Y.; Cramer, C. J.; Truhlar, D. G. The Interface of Electronic Structure and Dynamics for Reactions in Solution. *Int. J. Quantum Chem.* **1998**, *70*, 887−896.
(40) Castejon, H.; Wiberg, K. B. Solvent Effects on Methyl Transfer Reactions. 1. The Menshutkin Reaction. *J. Am. Chem. Soc.* **1999**, *121*, 2139−2146.
(41) Naka, K.; Sato, H.; Morita, A.; Hirata, F.; Kato, S. RISM-SCF Study of the Free-Energy Profile of the Menshutkin-Type Reaction NH3+CH3Cl→NH3CH3++Cl− in Aqueous Solution. *Theor. Chem. Acc.* **1999**, *102*, 165−169.
(42) Su, P.; Ying, F.; Wu, W.; Hiberty, P. C.; Shaik, S. The Menshutkin Reaction in the Gas Phase and in Aqueous Solution: A Valence Bond Study. *ChemPhysChem* **2007**, *8*, 2603−2614.
(43) Gao, J.; Xia, X. A Two-Dimensional Energy Surface for a Type II SN2 Reaction in Aqueous Solution. *J. Am. Chem. Soc.* **1993**, *115*, 9667−9675.
(44) Okamoto, K.; Fukui, S.; Shingu, H. Kinetic Studies of Bimolecular Nucleophilic Substitution. VI. Rates of the Menschtkin Reaction of Methyl Iodide with Methylamines and Ammonia in Aqueous Solutions. *Bull. Chem. Soc. Jpn.* **1967**, *40*, 1920−1925.
(45) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(46) Paulechka, E.; Kazakov, A. Efficient DLPNO-CCSD(T)-Based Estimation of Formation Enthalpies for C-, H-, O-, and N-Containing Closed-Shell Compounds Validated Against Critically Evaluated Experimental Data. *J. Phys. Chem. A* **2017**, *121*, 4379−4387.
(47) Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An Improved Linear Scaling Perturbative Triples Correction for the Domain Based Local Pair-Natural Orbital Based Singles and Doubles Coupled Cluster Method [DLPNO-CCSD(T)]. *J. Chem. Phys.* **2018**, *148*, No. 011101.
(48) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.
(49) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(50) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-Parameter, Range-Separated Hybrid, Generalized Gradient Approximation Density Functional with Nonlocal Correlation, Designed by a Survival-of-the-Fittest Strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904−9924.
(51) Chai, J. Da; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(52) Neese, F.; Hansen, A.; Liakos, D. G. Efficient and Accurate Approximations to the Local Coupled Cluster Singles Doubles Method Using a Truncated Pair Natural Orbital Basis. *J. Chem. Phys.* **2009**, *131*, 064103.
(53) Jorgensen, W. L. Free Energy Calculations: A Breakthrough for Modeling Organic Chemistry in Solution. *Acc. Chem. Res.* **1989**, *22*, 184−189.
(54) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(55) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2009**, *31*, 671−690.
(56) Yu, W.; He, X.; Vanommeslaeghe, K.; MacKerell, A. D. Extension of the CHARMM General Force Field to Sulfonyl-Containing Compounds and Its Utility in Biomolecular Simulations. *J. Comput. Chem.* **2012**, *33*, 2451−2468.
(57) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. *Gaussian 16,* Revision B.01; Gaussian Inc: Wallingford, CT,2016.
(58) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; et al. Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package. *Mol. Phys.* **2015**, *113*, 184−215.
(59) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931−967.
(60) Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. Towards an Order-N DFT Method. *Theor. Chem. Acc.* **1998**, *99*, 391−403.
(61) Baerends, E. J.; Ziegler, T.; Atkins, A. J.; Autschbach, J.; Bashford, D.; Baseggio, O.; Bérces, A.; Bickelhaupt, F. M.; Bo, C.; Boerritger, P. M.; et al. *ADF2017, SCM,* Theoretical Chemistry, Vrije Universiteit: Amsterdam, The Netherlands, 2017. Online at https://www.scm.com.
(62) Neese, F. Software Update: The ORCA Program System, Version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, No. e1327.
(63) Neese, F. The ORCA Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73−78.
(64) Pye, C. C.; Ziegler, T.; van Lenthe, E.; Louwen, J. N. An Implementation of the Conductor-like Screening Model of Solvation within the Amsterdam Density Functional Package — Part II. COSMO for Real Solvents. *Can. J. Chem.* **2009**, *87*, 790−797.
(65) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(66) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(67) Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. *Mol. Phys.* **1970**, *19*, 553−566.
(68) Simon, S.; Duran, M.; Dannenberg, J. J. How Does Basis Set Superposition Error Change the Potential Surfaces for Hydrogen-bonded Dimers? *J. Chem. Phys.* **1996**, *105*, 11024−11031.
(69) 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.
(70) Wood, R. H.; Yezdimer, E. M.; Sakane, S.; Barriocanal, J. A.; Doren, D. J. Free Energies of Solvation with Quantum Mechanical Interaction Energies from Classical Mechanical Simulations. *J. Chem. Phys.* **1999**, *110*, 1329−1337.
(71) Wood, R. H.; Liu, W.; Doren, D. J. Rapid Calculation of the Structures of Solutions with Ab Initio Interaction Potentials. *J. Phys. Chem. A* **2002**, *106*, 6689−6693.
(72) Iftimie, R.; Salahub, D.; Wei, D.; Schofield, J. Using a Classical Potential as an Efficient Importance Function for Sampling from an Ab Initio Potential. *J. Chem. Phys.* **2000**, *113*, 4852−4862.
(73) Iftimie, R.; Schofield, J. Separation of Quantum and Classical Behavior in Proton Transfer Reactions: Implications from Studies of Secondary Kinetic Isotope Effects. *Int. J. Quantum Chem.* **2003**, *91*, 404−413.
(74) Bandyopadhyay, P. 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*, 091102.
(75) Beierlein, F. R.; Michel, J.; Essex, J. W. A Simple QM/MM Approach for Capturing Polarization Effects in Protein−Ligand Binding Free Energy Calculations. *J. Phys. Chem. B* **2011**, *115*, 4911−4926.
(76) Bentzien, J.; Muller, R. P.; Florián, J.; Warshel, A. Hybrid Ab Initio Quantum Mechanics/Molecular Mechanics Calculations of

DOI: 10.1021/acs.jpca.9b03995
J. Phys. Chem. A 2019, 123, 5580−5589

Free Energy Surfaces for Enzymatic Reactions: The Nucleophilic Attack in Subtilisin. *J. Phys. Chem. B* **1998**, *102*, 2293−2301.
(77) Heimdal, J.; Ryde, U. Convergence of QM/MM Free-Energy Perturbations Based on Molecular-Mechanics or Semiempirical Simulations. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12592−12604.
(78) Jia, X.; Wang, M.; Shao, Y.; König, G.; Brooks, B. R.; Zhang, J. Z. H.; Mei, Y. Calculations of Solvation Free Energy through Energy Reweighting from Molecular Mechanics to Quantum Mechanics. *J. Chem. Theory Comput.* **2016**, *12*, 499−511.
(79) Kearns, F.; Warrensford, L.; Boresch, S.; Woodcock, H. The Good, the Bad, and the Ugly: "HiPen", a New Dataset for Validating (S)QM/MM Free Energy Simulations. *Molecules* **2019**, *24*, 681.
(80) Ho, J.; Shao, Y.; Kato, J. Do Better Quality Embedding Potentials Accelerate the Convergence of QM/MM Models? *Molecules* **2018**, *23*, 2466.
(81) Kish, L. *Survey Sampling;* Wiley: New York, 1965.
(82) Ryde, U. How Many Conformations Need to Be Sampled to Obtain Converged QM/MM Energies? The Curse of Exponential Averaging. *J. Chem. Theory Comput.* **2017**, *13*, 5745−5752.
(83) Boresch, S.; Woodcock, H. L. Convergence of Single-Step Free Energy Perturbation. *Mol. Phys.* **2017**, *115*, 1200−1213.
(84) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. Transition Path Sampling: Throwing Ropes over Rough Mountain Passes, in the Dark. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.
(85) Chandler, D. Statistical Mechanics of Isomerization Dynamics in Liquids and the Transition State Approximation. *J. Chem. Phys.* **1978**, *68*, 2959−2970.
(86) Dellago, C.; Bolhuis, P. G.; Geissler, P. L. Transition Path Sampling. *Adv. Chem. Phys.* **2003**, *123*, 1−78.
(87) Zhou, Y.; Pu, J. Reaction Path Force Matching: A New Strategy of Fitting Specific Reaction Parameters for Semiempirical Methods in Combined QM/MM Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 3038−3054.

DOI: 10.1021/acs.jpca.9b03995
J. Phys. Chem. A 2019, 123, 5580−5589