JCTC Journal of Chemical Theory and Computation

Article

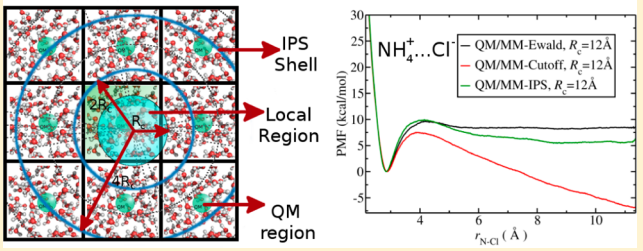
pubs.acs.org/JCTC

Isotropic Periodic Sum Treatment of Long-Range Electrostatic Interactions in Combined Quantum Mechanical and Molecular Mechanical Calculations

Pedro Ojeda-May and Jingzhi Pu*

Department of Chemistry and Chemical Biology, Indiana University-Purdue University Indianapolis, 402 N. Blackford St., LD326, Indianapolis, Indiana 46202, United States

ABSTRACT: The isotropic periodic sum (IPS) method was extended to describe long-range electrostatic interactions in combined quantum mechanical and molecular mechanical (QM/MM) calculations. The resulting method, designated QM/MM-IPS, was tested for two ion association processes and a model SN_2 reaction in aqueous solution. Potential of mean force (PMF) profiles and radial distribution functions computed from the QM/MM-IPS simulations were compared with those obtained by using the existing QM/MM-Ewald sum and cutoff (QM/MM-Cutoff) methods. In contrast to the QM/MM-Cutoff method, with which PMFs of ion separation tend to display a spurious linear drift, the QM/MM-IPS method successfully eliminates such artifacts, in excellent agreement with the QM/MM-Ewald results. The PMF obtained with the QM/MM-IPS method for the SN_2 reaction that transfers an NH_3 group between two chloride anions closely resembles that from the QM/MM-Ewald simulations. Compared with QM/MM-Ewald, the QM/MM-IPS method reduces the computational cost by 60−70% when a local region of 12 to 14 Å is used. These results suggest that the QM/MM-IPS method can be used as a reliable and efficient alternative to the QM/MM-Ewald method to incorporate long-range electrostatic effects in simulating solution-phase chemical reactions.



1. INTRODUCTION

Accurate and reliable treatment of long-range electrostatic interactions plays a crucial role in simulating chemical processes involving complex reactive systems in condensed phases and biological systems,[1−9] for which combined quantum mechanical and molecular mechanical (QM/MM)[10] potential energy functions are commonly used.

In the QM/MM approach, a subsystem is selected as the QM region and described by quantum mechanical wave functions (or electron density), whereas the rest of the system (the MM region), devoid of chemical changes, is treated by molecular mechanics. As the presence of the MM charges polarizes the QM wave function, accurate descriptions of the electrostatic interactions between the QM region and a large number of MM particles critically determine the chemical reactivity of the QM subsystem. In practical QM/MM simulations, however, such interactions are usually truncated beyond a cutoff distance to reduce the computational costs. Because electrostatic potential energy decays slowly over the distances r between the charged particles, the brute force truncation of the electrostatic interaction at a cutoff distance is a questionable approximation and has been related to a number of artifacts in classical simulations.[11−13] Despite the fact that numerical treatments exist to gradually attenuate electrostatic energies and the associated forces at the cutoff distance,[14−16] which may alleviate the problems of using such abrupt truncations, explicit inclusion of long-range electrostatic effects beyond the cutoff limit is highly desirable for reliable molecular simulations in general.

A host of techniques have been developed to incorporate long-range electrostatic interactions in the framework of molecular mechanical force fields by using the lattice summation methods. These include Ewald summation,[17] the particle-mesh Ewald (PME) method,[18,19] particle-particle-particle-mesh Ewald (P3M),[20] and the fast multipole methods (FMM).[21] Although other nonlattice alternatives such as the damped-shifted[15,16] methods or reaction field[22−24] calculations can be used to mimic or capture the long-range effects implicitly, the lattice summation approaches are widely considered the most accurate treatments in representing long-range electrostatic effects.

Despite intense efforts in developing long-range electrostatic treatments for classical MM force fields, such developments for combined QM/MM calculations have not received sufficient attention only until recently, partly owing to additional technical complexity when the QM wave function is involved. To date, long-range electrostatic effects can be included in QM/MM calculations primarily in two ways.

The first type of QM/MM long-range treatment is based on a finite system, typically defined by a spherical boundary[25] or a boundary of arbitrary shape.[26] For the finite representation of

Received: August 13, 2013
Published: November 26, 2013

the system, the long-range electrostatic contributions from the bulk solvent, often treated as a continuum dielectric, is implicitly introduced as a reaction field by making use of a boundary potential. One example using this kind of treatment is the generalized solvent boundary potential (GSBP) method,[26] originally developed by Im and Roux for classical simulations. In the GSBP method, the system is divided into an inner and outer region. While the charges in the inner region are represented explicitly and subject to dynamical movements, the outer-region charge distributions are fixed in the presence of implicit solvents described as a polarizable continuum. The electrostatic modifications of the inner region due to the presence of the outer region is described by a solvent boundary potential that mimics the effect of explicit inclusion of the long-range interactions between the two regions. The GSBP approach has been extended by Cui and co-workers for combining the self-consistent charge density functional tight-binding method with molecular mechanics (SCC-DFTB/MM),[7] as well as for ab initio QM/MM calculations,[27] and recently implemented and refined by Thiel and co-workers for semiempirical QM/MM calculations.[5,6] Although the QM/MM-GSBP treatments strike an impressive balance between computational efficiency and theoretical robustness, a known limitation associated with the method is that the electrostatic potential and elements of the reaction field matrix have to be obtained based on a predefined molecular configuration. Thus, there is some concern when the configurations during the simulations differ significantly from the one based on which the reaction field is determined; challenging cases include chemical reactions that are accompanied with large conformational changes.[28] The use of a fixed configuration for the outer region can result in a biased computation of static and dynamical properties of the system.[26] As obtaining the reaction fields requires extensive Poisson−Boltzmann calculations, the QM/MM-GSBP method can be computationally demanding for large-sized systems.

In contrast to the boundary potential methods that deal with a finite system, the second class of methods incorporates long-range electrostatics into QM/MM calculations by using periodically repeated systems on which lattice summation is performed. In the lattice sum QM/MM calculations, electrostatic interactions are collected explicitly through infinitely repeated periodic images of the primary cell system. A well-established representative of the lattice summation methods is Ewald summation,[17] which converts the conditionally convergent direct sum of Coulombic interactions between charged particles into two rapidly converging sums: a real-space sum made among rapidly decaying interactions of screened point charges surrounded by compensating charge distributions and a reciprocal-space sum carried out for canceling the effect of the screening charge distributions by making use of Fourier transform. Incorporation of the Ewald treatment into QM/MM calculations was originally demonstrated by the work of Gao et al.,[2] in which long-range Ewald potentials were introduced in the QM/MM interaction Hamiltonian through one-electron integrals such that they explicitly polarize the QM wave functions. A computationally more convenient formalism has recently been introduced by Nam et al.,[3] in which the use of approximated charge distributions for the QM region results in a (nearly)linear-scaling version of QM/MM-Ewald at semiempirical QM levels. Following a similar line of derivation, a separate implementation of treating long-range electrostatics in QM/MM has been achieved in conjunction with the particle-mesh Ewald (PME) technique, which in principle makes QM/MM-PME calculations computationally more efficient.[4] In addition to QM/MM-Ewald based on semiempirical QM methods, Ewald summation has also been developed for QM/MM calculations in the context of the SCC-DFTB method[8] and density functional theory (DFT) methods.[29−31]

Although Ewald summation has been commonly accepted as one of the most reliable methods in treating long-range electrostatics, there are several known limitations when it is applied to simulate condensed-phase systems. First, the application of lattice-based methods, including Ewald summation, requires the imposition of periodic boundary conditions (PBCs), which may introduce spurious periodicity when noncrystalline systems, such as biomolecules in infinite dilution, are simulated. The over exaggerated symmetry effect imposed by PBCs have been shown to give rise to various artifacts, such as unwanted long-range correlations[32] and spuriously induced anisotropy effects.[33] Such effects in Ewald calculations have been discussed for ion solvation/ion−ion interactions[34] and conformational distributions of biomolecules,[35,36] and have been reviewed elsewhere.[37] In addition, when a periodic lattice is used, the convergence of the Ewald sum requires charge neutralization of the primary unit cell. Although such conditions can be handled by introducing counterions in an ad hoc manner or by applying charge correction terms,[38] these procedures may introduce additional sources of error to Ewald calculations. For example, the nonzero charge effects on ion solvation using the Ewald-based approach have been thoroughly discussed by Hummer et al.[23,39] and by Hunenberger and co-workers;[40−42] the related analysis on protein systems has recently been reported by Lu and Cui.[43] Finally, we note that effects of imposed periodicity and treatment of charge neutralization have not been systematically studied for QM/MM-Ewald calculations.

As lattice methods explicitly sum electrostatic interactions over infinitely repeated periodic systems, the calculations involved are often computationally demanding. The original Ewald method is an $O(N^2)$ algorithm (N being the number of charged particles) and can be optimized to the order of $O(N \log N)$ by employing the PME technique.[18,19] The scaling behavior of PME is restricted by use of the fast Fourier transform, which intrinsically limits the application of PME to large-sized molecular systems, even on massively parallelized platforms.

Given the limitations of the lattice sum methods discussed above, it has become evident that the development of alternative treatments of long-range electrostatics that circumvent explicit summation over periodic lattice systems is highly desirable. Among a variety of the lattice-free long-range electrostatic treatments is the isotropic periodic sum (IPS) method.[44−53] The IPS method was originally developed by Wu and Brooks for nonpolar[44] and polar[46] homogeneous systems, later combined with the discrete fast Fourier transform (DFFT) for treating heterogeneous systems[45] and recently improved by others.[51,53] It has been found that the IPS method can produce results comparable to those of methods such as Ewald summation and the cutoff method for a variety of homogeneous and heterogeneous systems, including simple bulk liquids,[44] interfacial systems,[47,52] and lipids.[48] Additionally, the IPS method is advantageous over lattice sum methods in that (i) it does not introduce artificial symmetry by imposing lattice periodicity, (ii) it does not require charge neutralization, and (iii) it scales linearly with system size [thus an $O(N)$

algorithm]. Because of these reasons, the IPS method has emerged as a promising alternative to Ewald summation for treating long-range electrostatics in classical simulations.

In spite of its success in simulating physical processes employing MM force fields, the IPS method has not been developed for and applied to simulations based on combined QM/MM potentials. Such limitation blocks the opportunity of exploiting the IPS technique to describe long-range electrostatic interactions in simulating reactive systems in condensed phases and biological environments, for which QM/MM potentials are required. Further development of the IPS method to remove this limitation is the focus of the present work.

In this paper, we report an extension of the IPS method for treating long-range electrostatics in combined QM/MM calculations. We refer to the new method as the combined quantum mechanical and molecular mechanical isotropic periodic sum (QM/MM-IPS) method. To achieve a balance between accuracy and speed, which is especially important for obtaining statistically sufficient samplings in condensed-phase simulations, we chose semiempirical molecular orbital (MO) methods as the QM level in the present work. To incorporate long-range electrostatic contributions, we introduced an IPS potential energy term that modifies the ordinary QM/MM calculations. We demonstrated the effectiveness of the QM/MM-IPS method by applying it to simulations of two ion-association processes and a symmetric SN_2 reaction in aqueous solution.

The rest of the article is organized as follows. In Section 2, we first briefly review the theoretical backgrounds of the IPS method in the framework of classical mechanics and then present the detailed description of the QM/MM-IPS method. Computational details of the simulations to test the QM/MM-IPS method are described in Section 3. Section 4 contains results and discussion. Concluding remarks are given in Section 5.

2. THEORY

In this section, we first review the IPS method in the molecular mechanical framework, which is referred to as the classical IPS method in this paper. More details about the classical IPS method can be found in the original references of Wu and Brooks.[44−46] We then describe the formalism of the QM/MM-IPS method, where we introduce the IPS long-range electrostatic interactions through a term that modifies the effective Fock matrix in QM/MM calculations; such treatment resembles the procedure used by Nam et al.[3] in the QM/MM-Ewald method.

2.1. Classical IPS. In a classical IPS treatment based on the pairwise additive nonpolarizable molecular mechanical potentials, a local region is defined for each particle using a spherical cutoff distance R_c; only electrostatic interactions between particle center i and its nonbonded neighbors j found in the local region are calculated explicitly

$$\varepsilon_{ij}(r_{ij}) = \frac{q_i q_j}{r_{ij}} \quad (r_{ij} \le R_c) \tag{1}$$

where ε_{ij} denotes the Coulomb energy between atomic charges q_i and q_j, and r_{ij} refers to the distance between atom i and j. Note that on the right-hand side (r.h.s.) of eq 1, we omit the explicit dependency of ε_{ij} on q_i and q_j for brevity. In the IPS method, these explicitly calculated local-region Coulombic interactions are also referred to as the local interactions. By contrast, the long-range contributions, which account for electrostatic interactions beyond the local region, is treated implicitly in the IPS method by assuming that the rest of the system can be represented by images of the particles in the local region. Note that the images used in the IPS method are different from the lattice images under periodic boundary conditions. In PBC simulations, the positions of periodic lattice images are explicitly determined through predefined symmetry operations, whereas the images used in IPS only exist virtually such that they obey certain statistical distributions. In particular, the IPS method assumes an isotropic and periodic distribution of these virtual images (thus, they are also referred to as isotropic periodic images), with which the long-range electrostatic contribution in the IPS method is represented by interactions between particle center i's isotropic periodic images and its local nonbonded neighbor j.

Under the assumption that the system can be considered homogeneous and isotropic beyond the size of the local region R_c, the contribution from these virtual images is treated in a mean-field manner in the IPS method. To illustrate this assumption, one can define a heterogeneity scale, similar to the concept of homogeneity scale[45] coined by Wu et al. Heterogeneity scale for a given system is reached when attention is gradually zoomed in on a local region until the heterogeneous features of the local region make it no longer representative of the bulk system. A system can be considered homogeneous in a length scale greater than its heterogeneity scale. Thus, the underlying assumption of the IPS method is that the heterogeneity scale of a system is smaller than the size of the local region R_c. The introduction of the isotropic periodic images in the IPS method greatly reduces the complexity of computing long-range electrostatic interactions. Compared with lattice sum methods, which rely on summation over periodic lattice images under PBCs, the IPS method achieves its efficiency by replacing explicit pairwise interactions with average interactions for which analytic expressions exist through angular integration. On the other hand, the IPS method can also offer reliable descriptions of long-range electrostatic effects, provided that the homogeneous assumption of the method is satisfied.

In practice, to deposit the isotropic periodic images in three-dimensional space, Wu et al. further split the long-range interactions for $r > R_c$ into two parts: one for images that are randomly distributed and the other for images positioned along the axis that connects the two local charged particles. The sum of the contributions from random and axial images yields the long-range correction term in IPS. Combining the local-region and long-range contributions, the total electrostatic potential energy (E_i) involving atom i in the IPS treatment can be written as[45]

$$\begin{aligned} E_i &= \frac{1}{2}\sum_{j}^{\infty}\varepsilon_{ij}(r_{ij}) \approx \frac{1}{2}\sum_{r_{ij}\le R_c}\varepsilon_{ij}(r_{ij}) + \frac{1}{2}\sum_{r_{ij}\le R_c}\phi_{ij}(r_{ij},R_c) \\ &= \frac{1}{2}\sum_{r_{ij}\le R_c}[\varepsilon_{ij}(r_{ij}) + \phi_{ij}(r_{ij},R_c)] \\ &= \frac{1}{2}\sum_{r_{ij}\le R_c}\varepsilon_{ij}^{\mathrm{IPS}}(r_{ij},R_c) \end{aligned} \tag{2}$$

where the pairwise correction energy term $\phi_{ij}(r_{ij},R_c)$ represents the IPS long-range contribution to the electrostatic energy.

dx.doi.org/10.1021/ct400724d | J. Chem. Theory Comput. 2014, 10, 134−145

Note that a factor of 1/2 is included in eq 2 to avoid double counting of the pairwise electrostatic interactions when E_i is collected over all atomic centers. As a result, the sum of the pairwise IPS electrostatic potential energy, referred to as $\varepsilon_{ij}^{IPS}(r_{ij},R_c)$ in eq 2, runs only over local interacting pairs whose separation r_{ij} is shorter than or equal to the cutoff distance R_c, whereas the long-range contribution beyond R_c is incorporated through the IPS approximation. In general, the pairwise IPS electrostatic potential energy $[\varepsilon_{ij}^{IPS}(r_{ij},R_c)]$ can be written as

$$\varepsilon_{ij}^{IPS}(r_{ij},R_c) = \begin{cases} \varepsilon_{ij}(r_{ij}) + \phi_{ij}(r_{ij},R_c) & \text{if } r_{ij} \le R_c \\ 0 & \text{if } r_{ij} > R_c \end{cases} \quad (3)$$

The key idea of the IPS method lies in the choice of the functional form of $\phi_{ij}(r_{ij},R_c)$. This function allows us to think of the charges outside the local region as "copies" of the local charges placed in space according to the isotropic periodic distribution. As shown by Wu et al., closed-form expressions of the long-range electrostatic contribution $\phi_{ij}(r_{ij},R_c)$ exist for monopole−monopole interactions and can be implemented as a truncated polynomial function of the reduced distance r_{ij}/R_c without introducing noticeable errors:

$$\phi_{ij}(r_{ij}, R_c) = \frac{q_i q_j}{R_c}\left[\sum_{k=1}^{6} a_{2k}\left(\frac{r_{ij}}{R_c}\right)^{2k}\right] \quad (4)$$

where a's are parameters determined by numerical fitting. The values of the coefficients a's are listed in Table 1 for the so-called nonpolar IPS potential (IPSn).[44] Note that a similar functional form, called the polar IPS potential (IPSp), has been proposed for treating systems by considering screening effects and dipolar interactions.[46] Moreover, Wu et al. has also developed a discrete fast Fourier transform (DFFT) version of the IPS method, referred to as IPS-DFFT,[45] for treating heterogeneous systems when a large local region is preferred. In the present paper, we focus our discussion of the QM/MM-IPS method on the nonpolar IPS potential. Implementations and tests of the QM/MM-IPS method based on the polar IPS potential and the DFFT extension are left for future work. For simplicity, we refer to the nonpolar IPS potential as the IPS potential throughout the paper unless otherwise noted.

Table 1. Coefficients for Nonpolar IPS Potential (ref 44) as Given in Eq 4

	nonpolar IPS potential
a_2	0.3018373
a_4	0.0648044
a_6	0.0157855
a_8	0.0037972
a_{10}	0.0012149
a_{12}	−0.0000110

For convenience of deriving the QM/MM-IPS potential energy, we further define the charge-independent part of eq 4 as the reduced IPS pair-potential $\bar{\phi}_{ij}(r_{ij},R_c)$

$$\bar{\phi}_{ij}(r_{ij},R_c) = \frac{\phi(r_{ij}, R_c)}{q_i q_j} = \frac{1}{R_c}\left[\sum_{k=1}^{6} a_{2k}\left(\frac{r_{ij}}{R_c}\right)^{2k}\right] \quad (5)$$

The reduced IPS pair-potential for the nonpolar form of the IPS potential used in the present work with R_c = 10 Å is plotted in Figure 1.

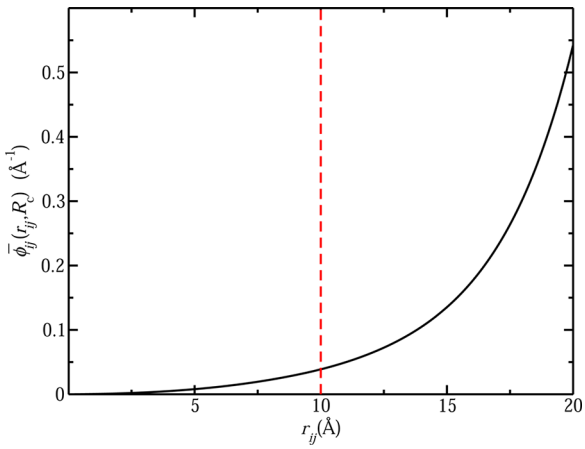


Figure 1. Plot of reduced IPS pair-potential $\bar{\phi}_{ij}(r_{ij},R_c)$ (defined in eq 5) for a cutoff distance of R_c = 10 Å. The polynomial part increases monotonically with the separation distance between atoms. The dashed line represents the boundary of the local region at $r = R_c$.

2.2. QM/MM-IPS. For a *finite* QM/MM system, the effective Hamiltonian ($\hat{\mathbf{H}}^{eff}$) can be written as a sum of three terms

$$\hat{\mathbf{H}}^{eff} = \hat{\mathbf{H}}_{QM} + \hat{\mathbf{H}}_{QM/MM} + \hat{\mathbf{H}}_{MM} \quad (6)$$

where $\hat{\mathbf{H}}_{QM}$ is the Hamiltonian for the QM subsystem in the gas phase, $\hat{\mathbf{H}}_{QM/MM}$ describes the electrostatic and van der Waals interactions between the QM and MM regions, and $\hat{\mathbf{H}}_{MM}$ contains bonded as well as nonbonded molecular mechanical interactions among MM particles. As the present work concerns electrostatic interactions, the discussion below only focuses on the electrostatic components of $\hat{\mathbf{H}}^{eff}$. For a QM/MM system whose QM region is treated by a molecular orbital (MO) theory method, the total electrostatic energy (E^{tot}) associated with $\hat{\mathbf{H}}^{eff}$ can be expreesed as

$$E^{tot} = E_{QM}(\psi) + E_{QM/MM}(\psi,\mathbf{q}) + E_{MM}(\mathbf{q}) \quad (7)$$

where ψ is the electronic wave function for the QM subsystem, and $\mathbf{q}$ represents a set of fixed point charges residing on the MM particles. With eq 7, we can attribute the total electrostatic energy of the QM/MM system to three contributions: the QM contribution $E_{QM}(\psi)$ due to the electronic wave function, the classical contribution $E_{MM}(\mathbf{q})$ due to the MM charges, and the QM/MM contribution $E_{QM/MM}(\psi,\mathbf{q})$ due to the interactions bewteen the wave function and the MM charges.

Similar to what has been shown by Nam et al.,[3] the total electrostatic interactions for an *infinite* QM/MM system can be represented as

$$E[\rho(\psi) + \mathbf{q},\rho(\psi) + \mathbf{q}] = E[\rho(\psi),\rho(\psi)] + E[\rho(\psi),\mathbf{q}] + E[\mathbf{q},\mathbf{q}] \quad (8)$$

where $\rho(\psi)$ denotes the charge distribution arising from the electronic wave function and the nuclear charges associated with the QM subsystem in the presence of a collection of MM charges $\mathbf{q}$. Note that for clarity, the square bracket notation [A,B] was introduced in eq 8 to explicitly specify the

interactions between two generalized charge distributions A and B. Once the key equations are obtained, the conventional function/variable notation will be restored. For brevity, $\rho(\psi)$ will be simply referred to as ρ below.

To describe long-range electrostatic interactions in a systematically improvable manner, one can split eq 8 into a local-region term and a long-range term that accounts for corrections needed to recover the electrostatics at full strength

$$E[\rho + \mathbf{q}, \rho + \mathbf{q}] = (E[\rho,\rho] - E^{\mathrm{LcR}}[\rho,\rho]) + E^{\mathrm{LcR}}[\rho,\rho] + (E[\rho,\mathbf{q}] - E^{\mathrm{LcR}}[\rho,\mathbf{q}]) + E^{\mathrm{LcR}}[\rho,\mathbf{q}] + E[\mathbf{q},\mathbf{q}] \quad (9)$$

where the superscript "LcR" is added to denote the local-region interactions bounded by the cutoff distance R_c. $E[\mathbf{q},\mathbf{q}]$ collects the electrostatic interactions among MM charges in an infinite system and thus can be treated separately by any existing classical long-range electrostatic methods. It should be emphasized that eq 9 is exact without making any approximation to eq 8. Direct application of eq 9 in QM/MM-IPS calculations, however, is often difficult because the computation of $E[\rho,\rho]$ and $E[\rho,\mathbf{q}]$ for an infinite system involving an arbitrary-shaped QM charge distribution ρ can be mathematically intricate and computationally demanding. In practice, one can approximate the QM charge distribution by multipole expansions and further simplifications can be made by truncating the expansion as a set of monopoles residing on the QM atoms. The particular monopole representation of the QM charge distribution we considered here is Mulliken charges,[54] which can be obtained from population analysis of the QM wave function. For semiempirical QM wave functions under the neglect of diatomic differential overlap (NDDO) approximation,[55] the Mulliken charge Q_α on the QM atom α is expressed as

$$Q_\alpha = Z_\alpha - \sum_{\mu\in\alpha} \rho_{\mu\mu} \quad (10)$$

where Z_α is the nuclear charge on a QM atom α, and $\rho_{\mu\mu}$ is the diagonal element of the electronic density matrix associated with the atomic orbital basis function μ that belongs to α.

In this way, we can obtain an approximate form of eq 9 as

$$E[\rho+\mathbf{q},\rho+\mathbf{q}] \approx (E[\mathbf{Q},\mathbf{Q}] - E^{\mathrm{LcR}}[\mathbf{Q},\mathbf{Q}]) + E^{\mathrm{LcR}}[\rho,\rho] + (E[\mathbf{Q},\mathbf{q}] - E^{\mathrm{LcR}}[\mathbf{Q},\mathbf{q}]) + E^{\mathrm{LcR}}[\rho,\mathbf{q}] + E[\mathbf{q},\mathbf{q}] \quad (11)$$

where $\mathbf{Q}$ represents a set of Mulliken charges localized on the QM atoms. The expression in eq 11 separates the total electrostatic interactions involving the QM region into local-region (LcR) contributions and the corresponding long-range corrections, as described by the two terms appeared in the parentheses. Note that although eqs 8 and 9 are exact, eq 11 is an approximated form because the realistic charge distribution is represented by multipole expansions truncated at monopoles (e.g., Mulliken charges); the underlying assumption that eq 11 can be used as a reasonable approximation to eq 9 is that the long-range correction terms in eq 11 are less sensitive to truncation of the charge distribution. As a remark, the long-range corrections needed in the QM/MM-IPS treatment are equal, in our case, to the IPS long-range contributions. Therefore, we can rewrite eq 11 as

$$E(\rho,\mathbf{q}) = \Delta E^{\mathrm{IPS}}_{\mathrm{QM}}(\mathbf{Q}) + E^{\mathrm{LcR}}_{\mathrm{QM}}(\rho) + \Delta E^{\mathrm{IPS}}_{\mathrm{QM/MM}}(\mathbf{Q},\mathbf{q}) + E^{\mathrm{LcR}}_{\mathrm{QM/MM}}(\rho,\mathbf{q}) + E^{\mathrm{IPS}}_{\mathrm{MM}}(\mathbf{q}) \quad (12)$$

where

$$\Delta E^{\mathrm{IPS}}_{\mathrm{QM}}(\mathbf{Q}) = E[\mathbf{Q},\mathbf{Q}] - E^{\mathrm{LcR}}[\mathbf{Q},\mathbf{Q}] \quad (13)$$

$$E^{\mathrm{LcR}}_{\mathrm{QM}}(\rho) = E^{\mathrm{LcR}}[\rho,\rho] \quad (14)$$

$$\Delta E^{\mathrm{IPS}}_{\mathrm{QM/MM}}(\mathbf{Q},\mathbf{q}) = E[\mathbf{Q},\mathbf{q}] - E^{\mathrm{LcR}}[\mathbf{Q},\mathbf{q}] \quad (15)$$

$$E^{\mathrm{LcR}}_{\mathrm{QM/MM}}(\rho,\mathbf{q}) = E^{\mathrm{LcR}}[\rho,\mathbf{q}] \quad (16)$$

$$E^{\mathrm{IPS}}_{\mathrm{MM}}(\mathbf{q}) = E(\mathbf{q},\mathbf{q}) \quad (17)$$

Note that starting from eq 12, we have simplified the notation by replacing the square-bracket notation with the conventional function/variable notation, such that the interacting pairs are no longer specified explicitly. For example, the total electrostatic interaction energy $E[\rho+\mathbf{q},\rho+\mathbf{q}]$ is now referred to as $E(\rho,\mathbf{q})$ because E depends on both ρ and $\mathbf{q}$; furthermore, we omit Mulliken charges $\mathbf{Q}$ from the argument list of a function if that function depends also on ρ because $\mathbf{Q}$ can be determined from ρ by eq 10. By regrouping the local-region and long-range IPS correction terms, we can rewrite eq 12 into an alternative form

$$E(\rho,\mathbf{q}) = E^{\mathrm{LcR}}(\rho,\mathbf{q}) + \Delta E^{\mathrm{IPS}}(\mathbf{Q},\mathbf{q}) + E^{\mathrm{IPS}}_{\mathrm{MM}}(\mathbf{q}) \quad (18)$$

where

$$E^{\mathrm{LcR}}(\rho,\mathbf{q}) = E^{\mathrm{LcR}}_{\mathrm{QM}}(\rho) + E^{\mathrm{LcR}}_{\mathrm{QM/MM}}(\rho,\mathbf{q}) \quad (19)$$

$$\Delta E^{\mathrm{IPS}}(\mathbf{Q},\mathbf{q}) = \Delta E^{\mathrm{IPS}}_{\mathrm{QM}}(\mathbf{Q}) + \Delta E^{\mathrm{IPS}}_{\mathrm{QM/MM}}(\mathbf{Q},\mathbf{q}) \quad (20)$$

The local-region (LcR) electrostatic contribution, as it is associated with a finite QM/MM system, can be computed according to eq 7, by collecting QM related terms

$$E^{\mathrm{LcR}}(\rho,\mathbf{q}) = E_{\mathrm{QM}}(\psi) + E^{\mathrm{LcR}}_{\mathrm{QM/MM}}(\psi,\mathbf{q}) \quad (21)$$

where ψ is the QM wave function polarized under the electric fields of MM charges. The LcR energy expressed in eq 21 is determined through Hartree−Fock self-consistent-field (SCF) MO calculations by solving the Roothaan-Hall equation[56]

$$\mathbf{F}^{\mathrm{LcR}}\mathbf{C}^{\mathrm{LcR}} = \mathbf{S}\mathbf{C}^{\mathrm{LcR}}\mathbf{E}^{\mathrm{LcR}} \quad (22)$$

where $\mathbf{F}^{\mathrm{LcR}}$, $\mathbf{C}^{\mathrm{LcR}}$, and $\mathbf{S}$ denotes the Fock operator, the molecular orbital coefficient matrix, and the overlap matrix, respectively; $\mathbf{E}^{\mathrm{LcR}}$ is the diagonal matrix containing molecular orbital energies; and the superscript "LcR" is used to remind the readers that when these matrices are constructed only the local-region electrostatic interactions are included.

The elements of the total Fock matrix ($\mathbf{F}^{\mathrm{tot}}$) that incorporates the IPS long-range contributions can be then written as

$$F^{\mathrm{tot}}_{\mu\nu} = F^{\mathrm{LcR}}_{\mu\nu} + \Delta F^{\mathrm{IPS}}_{\mu\nu} \quad (23)$$

where μ and ν refer to a pair of atomic orbital basis functions. These individual Fock matrix elements can be computed from the energy derivatives with respect to the corresponding density matrix elements $\rho_{\mu\nu}$

$$F_{\mu\nu}^{\mathrm{LcR}} = \frac{\delta E^{\mathrm{LcR}}(\rho,\mathbf{q})}{\delta\rho_{\mu\nu}} \tag{24}$$

and

$$\Delta F_{\mu\nu}^{\mathrm{IPS}} = \frac{\delta \Delta E^{\mathrm{IPS}}(\mathbf{Q},\mathbf{q})}{\delta\rho_{\mu\nu}} \tag{25}$$

Making use of the definition of Mulliken charges in eq 10, the required IPS corrections projected to Fock matrix elements on the right-hand side (r.h.s) of eq 23 can be computed as

$$\begin{aligned}\frac{\delta \Delta E^{\mathrm{IPS}}(\mathbf{Q},\mathbf{q})}{\delta\rho_{\mu\nu}} &= \sum_{\mu,\nu\in\alpha} \frac{\delta Q_\alpha}{\delta\rho_{\mu\nu}} \times \frac{\partial \Delta E^{\mathrm{IPS}}(\mathbf{Q},\mathbf{q})}{\partial Q_\alpha} \\ &= -\sum_{\mu,\nu\in\alpha} \delta_{\mu\nu} \times \frac{\partial \Delta E^{\mathrm{IPS}}(\mathbf{Q},\mathbf{q})}{\partial Q_\alpha}\end{aligned} \tag{26}$$

where $\delta_{\mu\nu}$ is Kronecker's delta function. The two components of the IPS contribution to the total electrostatic energy, i.e., $\Delta E^{\mathrm{IPS}}(\mathbf{Q},\mathbf{q})$ in eq 20, can be written explicitly as

$$\Delta E_{\mathrm{QM}}^{\mathrm{IPS}}(\mathbf{Q}) = \frac{1}{2}\sum_{\alpha}^{N_{\mathrm{QM}}}\sum_{\beta\neq\alpha}^{N_{\mathrm{QM}}} Q_\alpha Q_\beta \bar{\phi}_{\alpha\beta}(r_{\alpha\beta},R_c) \tag{27}$$

and

$$\Delta E_{\mathrm{QM/MM}}^{\mathrm{IPS}}(\mathbf{Q},\mathbf{q}) = \sum_{\alpha}^{N_{\mathrm{QM}}}\sum_{j}^{N_{\mathrm{MM}}} Q_\alpha q_j \bar{\phi}_{\alpha j}(r_{\alpha j},R_c) \tag{28}$$

where α, β run over the QM atoms and j over the MM atoms; the number of QM and MM atoms is denoted by N_{QM} and N_{MM}, respectively; and $\bar{\phi}_{kl}(r_{kl},R_c)$ is the reduced IPS pair-potential, defined in eq 5. The nuclear gradients associated with the IPS contributions in eqs 27 and 28 can be calculated straightforwardly with chain rules because they are ordinary polynomial functions of r.

2.3. Boundary Energy. As Wu et al. pointed out,[44] the pairwise IPS electrostatic potential energy $\varepsilon_{ij}^{\mathrm{IPS}}(r_{ij},R_c)$, defined by eq 3, is not a continuous function at the cutoff distance $r_{ij} = R_c$. As a result, a given charged particle would experience an abrupt energy change when it moves across the boundary of local-region sphere; the amount of this energy change is referred to as the boundary energy[44] (Figure 2a). Such energy discontinuity, if not handled properly, would cause ill-conditioned sampling of the configuration space when Monte Carlo (MC) sampling is used. The boundary effect, however, does not seem to cause any direct consequence in molecular dynamics (MD) simulations because the associated forces for integrating the Newton's equation remain continuous when the particle crosses the boundary of the local region; recall that in the IPS treatment the local-region and long-range forces cancel out exactly to generate zero net forces at the boundary[44] (see eq 16 of ref 44). Nevertheless, for MD simulations where potential energy information are explicitly used, such as alchemical or free energy perturbation simulations, extra care has to be taken if energy is not continuous. Therefore, it is desirable to maintain continuity for both energy and force in IPS calculations. Classically, this problem is solved by computing an average of the boundary energies in the system and adding the averaged boundary energy, as a fixed energy shift, to the total energy of the system. This procedure, however, cannot be directly applied to QM/MM-IPS calculations in which the boundary energy term is not a constant. Instead, boundary energies vary for each instantaneous configuration sampled along QM/MM-IPS MD simulations because the magnitude of the boundary energy term depends on Mulliken charges, which need to be updated each time when the self-consistent-field (SCF) calculation is performed for the QM/MM system.

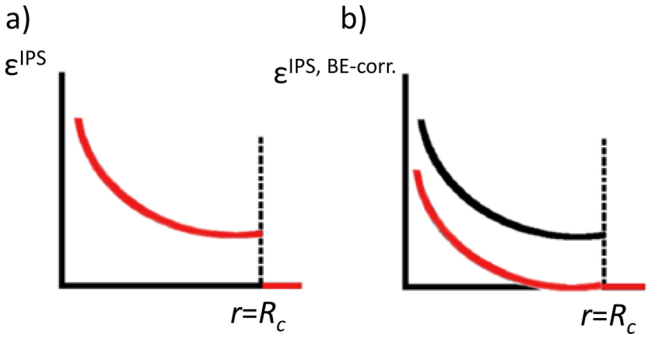


Figure 2. Schematic representation for the boundary energy treatment in the QM/MM-IPS method. (a) Total electrostatic interaction energy $\varepsilon^{\mathrm{IPS}}$ between a pair of charged particles in IPS, as described by eq 3, experiences discontinuity across the boundary of the local region at $r = R_c$ if no boundary energy correction is made. (b) For each pair of particles, the IPS contribution-included total electrostatic potential energy with boundary energy correction $\varepsilon^{\mathrm{IPS,BE\text{-}corr.}}$ (red line), is computed by subtracting the boundary energy, BE = ϕ^{IPS} ($r = R_c$) from the potential energy $\varepsilon^{\mathrm{IPS}}$ (r, R_c) (black line).

We addressed this issue by subtracting the instantaneous boundary energies, defined as the QM/MM-IPS potential energy computed at $r = R_c$ for each pair of particles, from the total pairwise QM/MM-IPS electrostatic potential energy (as illustrated in Figure 2b). As the boundary energy now is determined consistently through SCF calculations, both energy and forces are continuous functions across the boundary of the local region at $r = R_c$.

2.4. Treatment of IPS Interactions between MM Particles. For heterogeneous systems, the IPS potential between MM particles can be treated in a sophisticated manner beyond the expression in eq 4. One such treatment is to use the discrete fast Fourier transform (DFFT) technique as shown by Wu et al.[45] With the IPS-DFFT treatment, a local region as large as twice of the simulation box length can be used without adding substantial computational costs to IPS calculations. To test the present QM/MM-IPS algorithm, one possible combination is to treat QM/MM-IPS with the nonpolar IPS potential, while treating the MM/MM electrostatic interactions by the IPS-DFFT procedure for a more accurate representation of long-range electrostatic effects among MM particles. Such a combination, however, would introduce a few inconsistencies, including the use of different functional forms for IPS potentials and different sizes of local regions used for QM/MM and MM/MM interactions. If we use a small local region for QM/MM interactions and a larger local region for MM/MM interactions, it introduces an imbalanced long-range electrostatic treatment when QM particles are involved. If we adopt a very large cutoff distance for QM/MM, to be consistent with the IPS-DFFT convention in handling MM/MM interactions, the current QM/MM-IPS implementation alone is too computationally demanding without aid of DFFT. To avoid these inconsistencies, we decided to treat MM/MM-IPS interactions also by a simple nonpolar IPS potential. In particular, we

implemented the nonpolar IPS potential in eq 4 with parameters given in Table 1 for MM/MM-IPS interactions. Altogether, both QM/MM-IPS and MM/MM-IPS interactions in our development are treated in a consistent manner.

2.5. Nonbonded Interaction Treatment in QM/MM-IPS. The group-based scheme used to determine the nonbonded interaction list for the QM region, as used for QM/MM-Ewald and QM/MM-Cutoff simulations, seems to be problematic for QM/MM-IPS simulations. According to the group-based scheme, when an MM group is within the cutoff sphere of at least one QM group (based on the center of mass-distance criterion), all point charges on particles in that MM group would be included into the QM/MM electrostatic calculations to polarize the entire QM region, even when distances between this MM group and some other QM groups are greater than R_c. Applying this treatment in IPS calculations, however, represents a departure from eq 3, especially when a uniform R_c is used for the whole system. If the group-based scheme is employed in QM/MM-IPS calculations, there would be situations where the reduced IPS pair-potential, defined in eq 4, beyond the size of the local region size R_c, has to be used to compute the IPS long-range contributions. Because the reduced IPS pair-potential increases rapidly for $r > R_c$ (Figure 1), the use of it outside the local region may lead to numerical instabilities in QM/MM-IPS calculations. We solved this problem by modifying the group scheme, i.e., we neglect IPS contributions from any QM/MM interacting pairs whose separation is greater than R_c in QM/MM-IPS calculations. We need to emphasize that this treatment is consistent with a uniform definition of local region across the whole system. With this modified group-based scheme, stable QM/MM-IPS simulations can be achieved satisfactorily. An alternative treatment for the group-based cutoff scheme is to use a reflection of the reduced IPS-pair potential beyond R_c, which can be rationalized by a minimum image convention under the isotropic periodic assumption, i.e., once a particle leaves the local-region sphere ($r > R_c$), its minimum image re-enters into the primary system from the other side of sphere, leading to an effective interatomic distance of $2R_c - r$ (Wu, personal communication). Testing the reflected IPS pair-potential for QM/MM-IPS calculations is under way.

3. COMPUTATIONAL DETAILS

We implemented the QM/MM-IPS method in the MOPAC[57]-based semiempirical module QUANTUM of a modified version of CHARMM[58] (version c36a1). To test the QM/MM-IPS method, we first performed simulations for two solution-phase ion association reactions. The former examines association of ammonium and chloride ions ($[NH_4^+...Cl^-]$), and the latter describes the association of ammonium and metaphosphate ions ($[NH_4^+...PO_3^-]$). In both cases, the solutes NH_4^+, Cl^-, and PO_3^- molecules are treated quantum mechanically, and the solvent molecules were treated molecular mechanically. In particular, we used the semiempirical AM1[59] and the MNDO[60] models, for $[NH_4^+...Cl^-]$ and $[NH_4^+...PO_3^-]$ systems, respectively. In addition, we also tested the QM/MM-IPS method for a chemical reaction that involves bond breaking and forming. In particular, we tested the following SN_2 model reaction in solution

$$[ClNH_3]^+ + Cl^- \rightarrow Cl^- + [NH_3Cl]^+ \quad (29)$$

where we treated the solute ($[ClNH_3]^+ + Cl^-$) by the AM1 Hamiltonian and the solvent water molecules by MM. We chose this overall neutral reaction instead of the more canonical form for a SN_2 reaction such as $ClCH_3 + Cl^- \rightarrow Cl^- + CH_3Cl$ for two reasons. First, the charge neutrality in reaction described by eq 29 eliminates uncertainties related to net-charge correction schemes in QM/MM-Ewald simulations, making the comparison between QM/MM-Ewald and the other two methods straightforward. Second, the reaction of eq 29 goes through a transition state that has zero dipole moment (by symmetry), compared to a much more polar reactant complex species; different solvation effects along the reaction coordinate are expected to play an important role in determining the reaction free energy profiles, making the quantitative features of the free energy barriers sensitive to the treatment of long-range electrostatics between the solute and solvent. Schematic representations of the three QM/MM systems are shown in Figure 3, where the corresponding QM region is highlighted by a green circle. In all three cases, a cubic periodic box of 40 Å × 40 Å × 40 Å of modified TIP3P[61] water molecules was used to solvate the reactive systems. The SHAKE algorithm[62] was used to constrain the internal geometries of waters during the MD simulations.

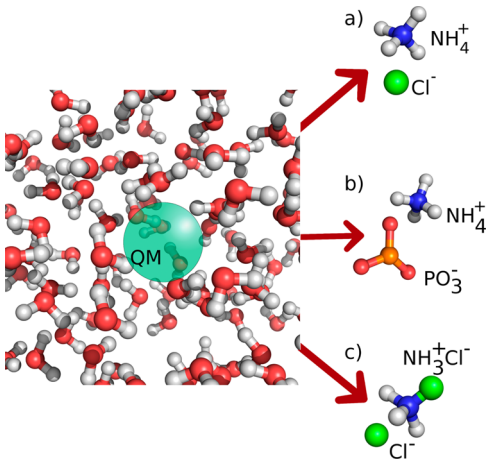


Figure 3. Schematic of the division between QM and MM regions. The QM region consists of (a) ammonium and chloride ions $[NH_4^+...Cl^-]$, (b) ammonium and metaphosphate ions $[NH_4^+...PO_3^-]$, and (c) a model SN_2 reaction $[ClNH_3]^+ + Cl^-$. The MM region is represented by the surrounding water molecules described by the modified TIP3P potential.

To benchmark the performance of the QM/MM-IPS method, we also carried out simulations with the QM/MM-Ewald and cutoff (QM/MM-Cutoff) methods. In all cases, we adopted periodic boundary conditions in the simulations. In the QM/MM-Ewald simulations, the κ parameter that represents the width of the Gaussian screening charge distributions was set to 0.34 $Å^{-1}$,[3] and for interactions among pure MM charges, PME was used for efficiency, in which the reciprocal space summations were performed on a 50 × 50 × 50 FFT grid, with maximally up to 5 k-vectors included in each Cartesian direction. For the QM/MM-Cutoff and QM/MM-IPS simulations, we used two cutoff distances at values of 12.0 and 14.0 Å, specified by the keyword "*cutnb*" in CHARMM, for the nonbonded list generation and for the size of the local region (R_c); the same *cutnb* parameters were also used to specify the cutoffs of the real space sum in QM/MM-Ewald

simulations, as well as for truncation of the van der Waals interactions, whose long-range contributions are neglected in the present work. In the QM/MM-Cutoff and QM/MM-IPS simulations, an energy switching function was applied to attenuate the electrostatic interactions between 10−12 or 12−14 Å in the two cutoff cases. Canonical ensembles at constant temperature and constant volume (NVT) were generated at 298 K by MD simulations. Langevin dynamics were employed to integrate equations of motion by using a time step of 1 fs, with friction coefficients on all atoms set to 50 ps^{-1}.

To determine potential energy of mean force (PMF)[63] free energy profiles of these solution-phase reactions, we employed the umbrella sampling technique.[64] For the $[NH_4^+...Cl^-]$ system, we chose the distance between the N and Cl atoms (r_{N-Cl}) as the reaction coordinate, whereas for the $[NH_4^+...PO_3^-]$ system, the reaction coordinate was chosen as the distance between the N and P atoms (r_{N-P}). A total of 37 windows were analyzed in the range of 2 Å < $r_{N-Cl(P)}$ < 11 Å. In the regions where the ion separations are small (2 Å< $r_{N-Cl(P)}$ < 3 Å), the force constants were chosen between 80−150 kcal/mol/Å^2, while for the intermediate separation regions (3 Å < $r_{N-Cl(P)}$ < 4 Å), the force constants were chosen in the range of 40−80 kcal/mol/Å^2. For larger distances of ion separation, the force constants were taken between 10−40 kcal/mol/Å^2.

As for the SN_2 reaction (eq 29), we constrained the Cl...Cl distance at 3.6 Å and the angle Cl−N−Cl at 180°. The reaction coordinate (X^R) was chosen as

$$X^R = r_{Cl-N} - r_{N-Cl} \quad (30)$$

A total of 25 windows were used in umbrella sampling simulations of this chemical reaction. A uniform force constant of 80 kcal/mol/Å^2 was applied on X^R to sample the system about the center of each reaction coordinate window.

With these schemes, we obtained umbrella sampling windows with properly overlapped probability histograms. The system for each window was first equilibrated by 30 ps QM/MM MD simulations and then further sampled by 50 ps simulations for data production. Finally, the PMFs were determined by the weighted histogram analysis method (WHAM).[65] The radial distribution functions (RDFs) of selected solute atoms and the solvent oxygen atom were computed and normalized with respect to the bulk density of each system. To obtain smooth RDFs with sufficient statistical samplings, we performed long simulations of 1 ns for each case. The number of solvent molecules coordinated to the central solute atom within its first solvation shell was calculated by integrating the bulk-density weighted RDF in spherical coordinates from r = 0 up to the radial position at which the first minimum of the RDF is located.

4. RESULTS AND DISCUSSION

4.1. Ammonium Chloride Association. For the ammonium/chloride system ($[NH_4^+...Cl^-]$) (Figure 3a), the PMF profiles simulated by the QM/MM-Ewald, QM/MM-Cutoff, and QM/MM-IPS methods are shown in Figure 4. The PMFs from the QM/MM-Cutoff simulations show linear drifts in the whole range of the reaction coordinate. Such linear drifts have been attributed to an unbalanced treatment of short-range and long-range electrostatic interactions in cutoff-based QM/MM simulations.[3] The free energy drift in the QM/MM-Cutoff simulated PMFs becomes less pronounced when a cutoff of a 14 Å is used compared with the results obtained by using a

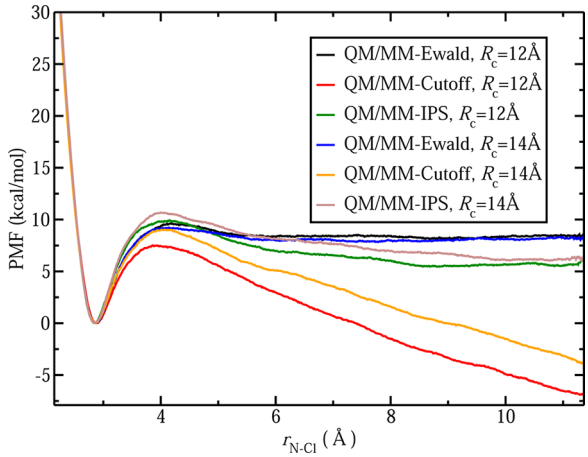


Figure 4. PMF profiles for $[NH_4^+...Cl^-]$ association by using the QM/MM-Ewald, QM/MM-Cutoff, and QM/MM-IPS methods. The QM/MM-IPS and QM/MM-Cutoff methods were computed at R_c = 12 and 14 Å; the same cutoffs were used for the real-space sum in the QM/MM-Ewald method. The PMFs by using the QM/MM-Cutoff method display linear drifts, whereas the QM/MM-IPS simulations do not generate the drifted PMFs, in agreement with the results of the QM/MM-Ewald method.

cutoff of 12 Å, in support of the electrostatic origin of this artifact. Both the QM/MM-IPS and QM/MM-Ewald methods incorporate long-range electrostatic effects, and therefore, they do not suffer from this artificial free energy drift as expected. The resulting PMFs in the QM/MM-IPS simulations are leveled off at large ion-separation distances, similar to the results obtained by using the QM/MM-Ewald method (Figure 4). It is a striking observation that although the QM/MM-IPS and QM/MM-Ewald methods describe long-range electrostatic interactions by following two very different philosophies, they seem to be equally effective in removing the electrostatic imbalance in the QM/MM-Cutoff simulations.

Next we discuss the quantitative features of the simulated PMFs. Free energy barriers of 8−10 kcal/mol were found around r_{N-Cl} = 4 Å in the PMFs simulated by all three methods. We note that this barrier was previously reported by Nam et al.[3] as ∼4 kcal/mol, who employed the QM/MM-Ewald method in constant temperature and constant pressure (NPT) simulations. We believe that the main difference in these QM/MM-Ewald simulations is the use of different ensembles (NVT in our simulations vs NPT in Nam et al.). Nevertheless, the quantitative features of PMFs obtained from our simulations by using two very different long-range electrostatic treatments, i.e., QM/MM-Ewald and QM/MM-IPS, based on the same setup seem to agree with each other very well (see below). This suggests that the simulations we conducted to evaluate the performance of the QM/MM-IPS method are self-contained. Therefore, we took our own QM/MM-Ewald simulations as the benchmark. The present QM/MM-Ewald simulations yield a barrier height of 9.6 and 9.2 kcal/mol when the cutoff value of 12 and 14 Å is used for the real-space sum. For distances greater than 5.5 Å, the PMFs from the QM/MM-Ewald simulations reach a saturation value of ∼8.5 kcal/mol (the saturation value reported by Nam et al.[3] is ∼0 kcal/mol). The free energy barrier heights identified from the QM/MM-IPS PMF profiles are 9.9 and 10.6 kcal/mol for the cutoff distance of 12 and 14 Å, respectively, only slightly higher than those obtained from the QM/MM-Ewald simulations. The saturation

dx.doi.org/10.1021/ct400724d | J. Chem. Theory Comput. 2014, 10, 134−145

values on the QM/MM-IPS PMFs are 5.6 and 6.2 kcal/mol for the cutoff value of 12 and 14 Å, respectively. Additionally, the free energy barrier heights for the QM/MM-Cutoff method are 7.4 and 9.0 kcal/mol for the cutoff distance of 12 and 14 Å, respectively. In the region of $r_{N\text{-}Cl} < 3.5$ Å, which is approximately centered about the reactant complex, all three methods produce similar PMFs.

The PMFs obtained from the QM/MM-IPS method at the two cutoff values are similar, suggesting that a size of 12 to 14 Å for the local region is sufficient to obtain converged QM/MM-IPS results, at least for the simple solution-phase reactions tested here. Therefore, further the structural analyses are focused only on one cutoff (see below).

The radial distribution functions (RDFs) of the Cl−O pair in the $[NH_4^+...Cl^-]$ system at the reactant complex state (r_{N-Cl} = 2.85 Å) obtained from the QM/MM-Ewald, QM/MM-Cutoff, and QM/MM-IPS simulations are compared in Figure 5. In this comparison, only RDF results from simulations using a cutoff distance of R_c = 12 Å are shown because the PMF results presented above suggest that the calculations converge reasonably well at this cutoff distance. We observed that the RDF obtained by the QM/MM-IPS method is similar to that of the QM/MM-Ewald method for the distance range of 2−10 Å. Compared with the two long-range-included QM/MM methods, the QM/MM-Cutoff method yields an RDF that contains a higher and narrower first peak and a more structured RDF at large distances. Based on the same ensembles (R_c = 12 Å; r_{N-Cl} = 2.85 Å), the averaged solvent coordination numbers within the first solvation shell of the solute Cl atom were computed for all three simulation methods (Table 2). The QM/MM-IPS method gives a first solvation shell value of 5.94 for the Cl−O pair, in close agreement with the value of 5.85 given by the QM/MM-Ewald method, whereas the QM/MM-Cutoff method gives a solvent coordination number of 5.99, which is the highest among the results obtained by the three methods. Taken together, these results suggest that the QM/MM-IPS method successfully reproduces solvation structural features of the QM/MM-Ewald method for the $[NH_4^+...Cl^-]$ system.

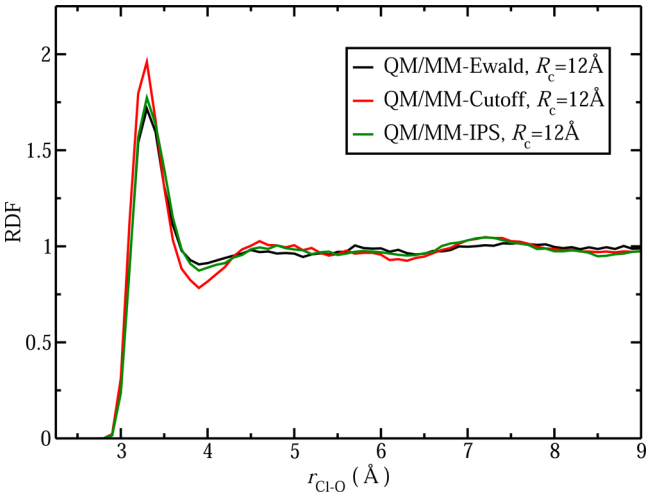


Figure 5. Cl−O RDFs in the $[NH_4^+...Cl^-]$ system simulated by the QM/MM-Ewald, QM/MM-Cutoff, and QM/MM-IPS methods at the reactant state ($r_{N\text{-}Cl}$ = 2.85 Å) for cutoff distance of the local region R_c = 12 Å (for the QM/MM-Ewald method, this cutoff was used in the real-space sum). RDF from the QM/MM-IPS simulation is comparable to that from the QM/MM-Ewald simulation.

4.2. Ammonium Metaphosphate Association. For the ammonium/metaphosphate system ($[NH_4^+...PO_3^-]$) (Figure 3b), the PMF profiles for the QM/MM-Ewald, QM/MM-Cutoff, and the QM/MM-IPS are shown in Figure 6. It should be noted that the PMF profiles were shifted so that the PMFs at r_{N-P} = 4 Å were set to zero free energy in all cases to make a consistent comparison. For this system, as opposed to the $[NH_4^+...Cl^-]$ system, no detectable free energy barriers were found with the QM/MM-Ewald method; the resulting PMFs are rather flat. The PMF of the QM/MM-Ewald reaches a constant value of −4.7 and −5.7 kcal/mol at distances >7 Å for

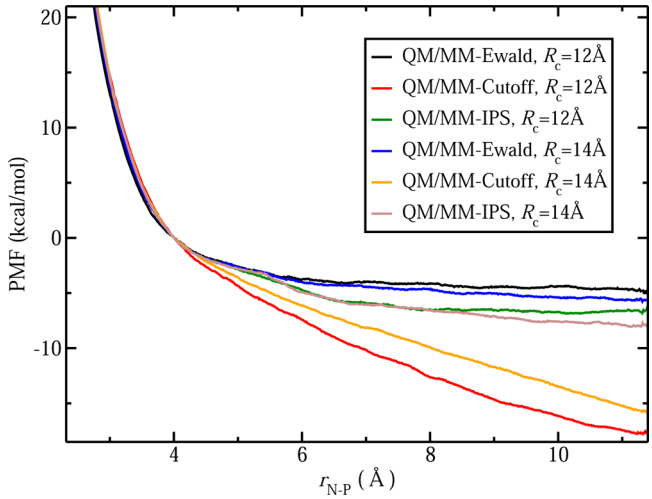


Figure 6. PMF profiles for $[NH_4^+...PO_3^-]$ association by using the QM/MM-Ewald, QM/MM-Cutoff, and QM/MM-IPS methods, evaluated at R_c = 12 and 14 Å (for the QM/MM-Ewald method, these cutoffs were used in the real-space sum). The PMF from the QM/MM-IPS method converges at these cutoff values. In agreement with the QM/MM-Ewald method, the QM/MM-IPS method treats long-range electrostatics in a balanced way; thus, the corresponding PMFs do not display the linear drifts seen in the QM/MM-Cutoff simulations.

Table 2. First Solvation Shell Computed for $[NH_4^+...Cl^-]$ and $[NH_4^+...PO_3^-]$ Systems with a Cutoff Distance of 12 Å

$[NH_4^+...Cl^-]$			
reference pairs	QM/MM-Ewald	QM/MM-Cutoff	QM/MM-IPS
Cl−O[a]	5.85	5.99	5.94
$[NH_4^+...PO_3^-]$			
reference pairs	QM/MM-Ewald	QM/MM-Cutoff	QM/MM-IPS
P−O[b]	14.62	14.75	14.59

[a]N−Cl distance is 2.85 Å. [b]N−P distance is 4.00 Å.

dx.doi.org/10.1021/ct400724d | J. Chem. Theory Comput. 2014, 10, 134−145

the real-space sum cutoff distance of 12 and 14 Å, respectively. The PMFs from the QM/MM-Cutoff method display a linear drift toward negative free energy values, similar to what have been observed in the QM/MM-Cutoff simulations for the $[NH_4^+...Cl^-]$ system. The PMFs generated by the QM/MM-IPS method level off at large distance with a saturation value of −6.7 and −8.0 kcal/mol for the cutoff distance of 12 and 14 Å, respectively, in good agreement with the QM/MM-Ewald results. This agreement, again, suggests that the long-range electrostatic interactions are treated satisfactorily by using the QM/MM-IPS method, whose quality is comparable to that of the QM/MM-Ewald method.

For the $[NH_4^+...PO_3^-]$ system, the PMFs obtained with the QM/MM-IPS method at the two cutoff values are similar, showing the same convergence behavior seen before in the $[NH_4^+...Cl^-]$ case; these results confirm that a local region of 12 to 14 Å is sufficient for the QM/MM-IPS method to describe long-range electrostatic effects reliably for the two ion-association reactions tested here.

The RDFs for the P−O pair in the $[NH_4^+...PO_3^-]$ system for the QM/MM-Ewald, QM/MM-Cutoff, and QM/MM-IPS methods are shown in Figure 7. We used a cutoff distance of

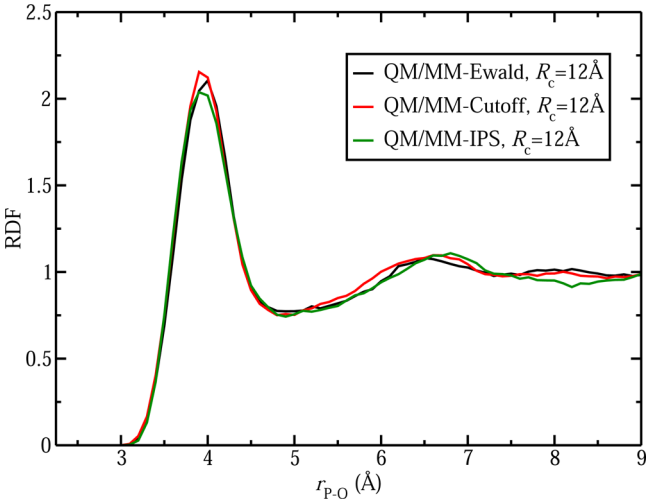


Figure 7. P−O RDFs in the $[NH_4^+...PO_3^-]$ system simulated by the QM/MM-Ewald, QM/MM-IPS, and QM/MM-Cutoff methods at the umbrella window $r_{N\text{-}P}$ = 4.00 Å. The cutoff distance was R_c = 12 Å in all cases (for the QM/MM-Ewald method, this cutoff was used in the real-space sum). RDF from the QM/MM-IPS simulation is comparable to that from the QM/MM-Ewald simulation.

R_c = 12 Å in all cases, as this cutoff seems to give reasonably converged results. The RDFs were computed for the umbrella window centered at r_{N-P} = 4.0 Å. The RDF obtained by using the QM/MM-IPS method is very similar to that from the QM/MM-Ewald simulation. Coordination numbers in the first solvation shell around the central P atom are also given in Table 2. Integration to the minimum of the first peak in the RDF from the QM/MM-IPS simulations yields a solvent coordination number of 14.59, in excellent agreement with the value of 14.62 based on the QM/MM-Ewald simulations, whereas the QM/MM-Cutoff simulations give a coordination number of 14.75.

4.3. Symmetric SN_2 Reaction between Chloride Anions. For the SN_2 reaction that symmetrically transfers an NH_3 group between two Cl^- anions (eq 29; Figure 3c), the PMF profiles obtained from the QM/MM-Ewald, QM/MM-Cutoff, and QM/MM-IPS simulations are shown in Figure 8. It is known that long-range interactions play an important role in

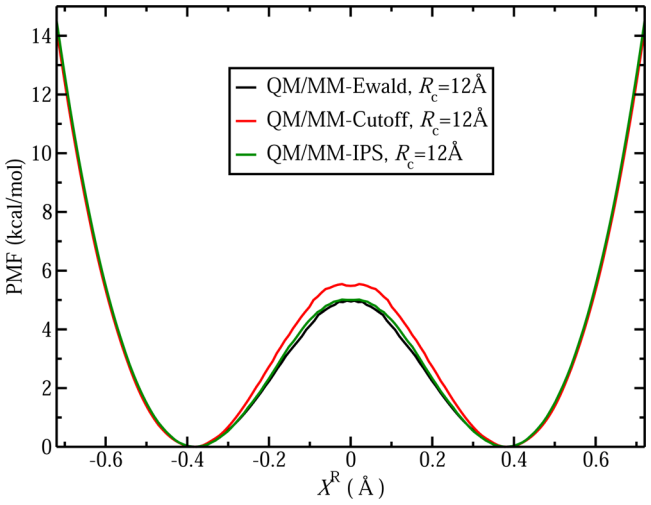


Figure 8. PMF profiles for the model SN_2 reaction $[ClNH_3^+ + Cl^-]$ by using the QM/MM-Ewald, QM/MM-Cutoff, and QM/MM-IPS methods, evaluated at R_c = 12 Å (for the QM/MM-Ewald method, this cutoff was used in the real-space sum). The QM/MM-IPS and QM/MM-Ewald method agree with each other by generating essentially same PMF free energy profiles, whereas the QM/MM-Cutoff method in comparison overestimates the free energy barrier height by 0.6 kcal/mol due to an unbalanced treatment of the local-region and long-range electrostatic solvation effects.

condensed-phase chemical reactions. Both the QM/MM-Ewald and QM/MM-IPS methods include treatments of long-range electrostatics, whereas the QM/MM-Cutoff method neglects such contributions beyond the cutoff distance. We found a significantly higher free energy barrier (0.6 kcal/mol) in the PMF from the QM/MM-Cutoff method compared to that from the QM/MM-IPS and QM/MM-Ewald simulations. This difference is likely caused by the treatment of long-range interactions beyond the cutoff sphere. The PMF obtained with the QM/MM-IPS method shows good agreement with the PMF from the QM/MM-Ewald simulations, suggesting that long-range electrostatics are handled properly in both methods.

Accurate treatments of long-range electrostatic interactions are crucial in modeling the SN_2 reaction we examined here because the reaction involves a reverse of the total dipole moment from the reactant to the product state. If symmetry is imposed, the transition state of this model SN_2 reaction corresponds to a state where the system's dipole moment is annihilated. Following this line of reasoning, the solvation effects are expected to change significantly along the reaction coordinate when the system evolves from the reactant (a more dipolar complex which is solvated better by polar solvents) to the transition state (less solvated because of the reduced charge delocalization and the lack of net dipole moment). Hence, different treatments of long-range electrostatic interactions between the solute and solvent, which describe this differential solvation effect to different extents, will likely lead to distinctions in the quantitative features of the free energy profile. When long-range electrostatic effects are neglected, one may expect that the absence of electrostatic solvation from distant solvent molecules would destabilize the polar reactant complex to a greater extent than it does to the transition state, resulting in a reduced free energy barrier. This predicted trend,

dx.doi.org/10.1021/ct400724d | J. Chem. Theory Comput. 2014, 10, 134−145

Table 3. Average CPU Time Over 37 Umbrella Sampling Windows of 80 ps QM/MM MD Simulations (time in h)

$[NH_4^+...Cl^-]$			
cutoff radius (Å)	QM/MM-Ewald	QM/MM-Cutoff	QM/MM-IPS
12	17.70	3.57	5.15
14	20.30	5.46	8.25
$[NH_4^+...PO_3^-]$			
cutoff radius (Å)	QM/MM-Ewald	QM/MM-Cutoff	QM/MM-IPS
12	19.16	3.75	5.34
14	20.75	5.74	8.47
$[ClNH_3]^+ + Cl^-$			
cutoff radius (Å)	QM/MM-Ewald	QM/MM-Cutoff	QM/MM-IPS
12	18.21	3.52	5.07

however, is opposite to the results obtained from our numerical simulations. The QM/MM-Cutoff method seems to overestimate the free energy barrier height, compared to the two long-range QM/MM methods. This observation could be attributed to possible artifacts in the QM/MM-Cutoff simulation, in which the lack of balanced long-range electrostatics may lead to over polarizations of the solute QM wave function by local solvent molecules, leading to spurious solvent stabilizations for the polar reactant species relative to the transition state (thus an elevated barrier). Interestingly, for the two ion separation processes we studied here, the finding that the free energy profiles produced by the QM/MM-Cutoff method always drift to the lower end of the free energy spectrum seems to imply that ions gain spurious stabilization when they are isolated, in accordance with this postulate. Other than that, both the QM/MM-Ewald and QM/MM-IPS methods, although with long-range electrostatic contributions included quite differently, agree with each other by yielding essentially the same free energy profiles for this SN_2 reaction.

4.4. Computational Efficiency. To evaluate the efficiency of the algorithm, we benchmarked the computational costs of the three simulation methods on a single 2 GHz quad-core Intel Xeon processor based on a single window of umbrella sampling QM/MM simulation. The average CPU times for umbrella sampling windows of 80 ps of the QM/MM-Ewald, QM/MM-Cutoff, and QM/MM-IPS simulations are given in Table 3. The average time (including the time spent on computing MM/MM interactions) to perform a single window of umbrella sampling with the QM/MM-IPS method (R_c = 12 Å) is only 1.4 times of that with the QM/MM-Cutoff method but 3.6 times less than that with the QM/MM Ewald method. For the cutoff distance of 14 Å, the QM/MM-IPS method is 1.4 times slower than the QM/MM-Cutoff method but 2.4 times faster than the QM/MM-Ewald method. On average, with the QM/MM-IPS method, the computational cost associated with treating long-range electrostatic interactions is greatly reduced by 60−70% compared to the QM/MM-Ewald method, yet one obtains similar results comparable to QM/MM-Ewald.

5. CONCLUDING REMARKS

In this work, we have extended the IPS method for treating long-range electrostatics in combined QM/MM calculations. To achieve an efficient and numerically stable algorithm, three important issues are addressed in this extension: (a) incorporation of the IPS long-range contributions to QM/MM calculations based on a Mulliken charge representation of the QM charge distribution, (b) self-consistent determination of the QM/MM-IPS boundary energies for each QM/MM interacting pairs, and (c) treatment of nonbonded interactions in QM/MM-IPS by using a modified group-based scheme. Tests on solution-phase ion association and SN_2 reactions show that satisfactory agreements on various condensed-phase properties, including PMF profiles, RDFs, and solvent coordination numbers, can be obtained between the QM/MM-IPS method and benchmark QM/MM-Ewald method. Finally, we show that the QM/MM-IPS method is approximately three times more efficient than the QM/MM-Ewald method for these simple solution-phase QM/MM systems. Because of the linear scaling behavior of the QM/MM-IPS method, greater gain on efficiency relative to the QM/MM-Ewlad method is expected when system size increases. Future work on the present development may include benchmarking the QM/MM-IPS method for ion solvation processes and testing it for more heterogeneous systems such as enzymatic reactions, for which a DFFT extension of the current algorithm is highly desirable.

AUTHOR INFORMATION

Corresponding Author

*E-mail: jpu@iupui.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Xiongwu Wu and Dr. Bernard Brooks for insightful discussions. This work was supported by a start-up grant from Indiana University-Purdue University Indianapolis (IUPUI). The computing time was provided by the High-Performance Computing Cluster, funded by the School of Science at IUPUI, and by the BigRed High Performance Computing facilities at Indiana University, funded by the National Science Foundation (NSF).

REFERENCES

(1) Ceperley, D. In NRCC Proceedings 9, Lawrence Berkeley National Laboratory: Berkeley, CA, 1980.
(2) Gao, J.; Alhambra, C. *J. Chem. Phys.* **1997**, *107*, 1212.
(3) Nam, K.; Gao, J.; York, D. M. *J. Chem. Theory Comput.* **2005**, *1*, 2.
(4) Walker, R. C.; Crowley, M. F.; Case, D. A. *J. Comput. Chem.* **2007**, *29*, 1019.
(5) Benighaus, T.; Thiel, W. *J. Chem. Theory Comput.* **2008**, *4*, 1600.
(6) Benighaus, T.; Thiel, W. *J. Chem. Theory Comput.* **2011**, *7*, 238.
(7) Schaefer, P.; Riccardi, D.; Cui, Q. *J. Chem. Phys.* **2005**, *123*, 014905.
(8) Riccardi, D.; Schaefer, P.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 17715.
(9) Kuwajima, S.; Warshel, A. *J. Chem. Phys.* **1988**, *89*, 3751.

(10) Gao, J., Thompson, M. A., Eds.; *Combined Quantum Mechanical and Molecular Mechanical Methods*; ACS Symposium Series 712; American Chemical Society: Washington, DC, 1998.
(11) Bader, J. S.; Chandler, D. *J. Phys. Chem.* **1992**, *96*, 6423.
(12) Smith, P. E.; Pettitt, B. M. *J. Chem. Phys.* **1991**, *95*, 8430.
(13) Brooks, C. L., III; Pettitt, B. M.; Karplus, M. *J. Chem. Phys.* **1985**, *83*, 5897.
(14) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667.
(15) Fennell, C. J.; Gezelter, J. D. *J. Chem. Phys.* **2006**, *124*, 234104.
(16) Wolf, D.; Keblinski, P.; Phillpot, S. R.; Eggebrecht, J. *J. Chem. Phys.* **1999**, *110*, 8254.
(17) Ewald, P. P. *Ann. Phys. (Leipzig)* **1921**, *369*, 253.
(18) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(19) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(20) Eastwood, J. W.; Hockney, R. W.; Lawrence, D. N. *Comput. Phys. Commun.* **1980**, *19*, 215.
(21) Greengard, L. *Science* **1994**, *265*, 909.
(22) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451.
(23) Hummer, G.; Pratt, L. R.; Garcia, A. E. *J. Chem. Phys.* **1997**, *107*, 9275.
(24) King, G.; Warshel, A. *J. Chem. Phys.* **1989**, *91*, 3647.
(25) Beglov, D.; Roux, B. *J. Chem. Phys.* **1994**, *100*, 9050.
(26) Im, W.; Bernèche, S.; Roux, B. *J. Chem. Phys.* **2001**, *114*, 2924.
(27) Zienau, J.; Cui, Q. *J. Phys. Chem. B* **2012**, *116*, 12522.
(28) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H.; Ghosh, N.; Prat-Resina, X.; Konig, P.; Li, G.; Xu, D.; Guo, H.; Elstner, M.; Cui, Q. *J. Phys. Chem. B* **2006**, *110*, 6458.
(29) Yarne, D. A.; Tuckerman, M. E.; Martyna, G. J. *J. Chem. Phys.* **2001**, *115*, 3531.
(30) Laino, T.; Mohamed, F.; Laio, A.; Parrinello, M. *J. Chem. Theory Comput.* **2005**, *1*, 1176.
(31) Laino, T.; Mohamed, F.; Laio, A.; Parrinello, M. *J. Chem. Theory Comput.* **2006**, *2*, 1370.
(32) Luty, B. A.; Tironi, I. G.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *103*, 3014.
(33) Kurtović, Z.; Marchi, M.; Chandler, D. *Mol. Phys.* **1993**, *78*, 1155.
(34) Hünenberger, P. H.; McCammon, J. A. *J. Chem. Phys.* **1999**, *110*, 1856.
(35) Weber, W.; Hünenberger, P. H.; McCammon, J. A. *J. Phys. Chem. B* **2000**, *104*, 3668.
(36) Hünenberger, P. H.; McCammon, J. A. *Biophys. Chem.* **1999**, *78*, 69.
(37) Tobias, D. J. *Curr. Opin. Struct. Biol.* **2001**, *11*, 253.
(38) Bogusz, S.; Cheatham, T. E., III; Brooks, B. R. *J. Chem. Phys.* **1998**, *108*, 7070.
(39) Hummer, G.; Pratt, L. R.; Garcia, A. E. *J. Phys. Chem.* **1996**, *100*, 1206.
(40) Kastenholz, M. A.; Hünenberger, P. H. *J. Chem. Phys.* **2006**, *124*, 124106.
(41) Kastenholz, M. A.; Hünenberger, P. H. *J. Chem. Phys.* **2006**, *124*, 224501.
(42) Reif, M. M.; Hünenberger, P. H.; Oostenbrink, C. *J. Chem. Theory Comput.* **2012**, *8*, 3705.
(43) Lu, X.; Cui, Q. *J. Phys. Chem. B* **2013**, *117*, 2005.
(44) Wu, X.; Brooks, B. R. *J. Chem. Phys.* **2005**, *122*, 044107.
(45) Wu, X.; Brooks, B. R. *J. Chem. Phys.* **2008**, *129*, 154115.
(46) Wu, X.; Brooks, B. R. *J. Chem. Phys.* **2009**, *131*, 024107.
(47) Klauda, J. B.; Wu, X.; Pastor, R. W.; Brooks, B. R. *J. Phys. Chem. B* **2007**, *111*, 4393.
(48) Venable, R. M.; Chen, L. E.; Pastor, R. W. *J. Phys. Chem. B* **2009**, *113*, 5855.
(49) Takahashi, K.; Yasuoka, K.; Narumi, T. *J. Chem. Phys.* **2007**, *127*, 114511.
(50) Takahashi, K.; Narumi, T.; Yasuoka, K. *J. Chem. Phys.* **2010**, *133*, 014109.
(51) Takahashi, K.; Narumi, T.; Yasuoka, K. *J. Chem. Phys.* **2011**, *135*, 174108.
(52) Takahashi, K.; Narumi, T.; Yasuoka, K. *J. Chem. Phys.* **2011**, *134*, 174112.
(53) Takahashi, K.; Narumi, T.; Suh, D.; Yasuoka, K. *J. Chem. Theory Comput.* **2012**, *8*, 4503.
(54) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.
(55) Pople, J. A.; Santry, D. P.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, S129.
(56) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.
(57) Stewart, J. J. P. *Quant. Chem. Prog. Exch.* **1990**, *10*, 86.
(58) Brooks, B. R.; C. L. Brooks, I.; A. D. Mackerell, J.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545.
(59) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(60) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899.
(61) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
(62) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327.
(63) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300.
(64) Torrier, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, *28*, 578.
(65) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011.