JCTC Journal of Chemical Theory and Computation

An Efficient Linear-Scaling Ewald Method for Long-Range Electrostatic Interactions in Combined QM/MM Calculations

Kwangho Nam, Jiali Gao, and Darrin M. York*

Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455-0431

Abstract: A method is presented for the efficient evaluation of long-range electrostatic forces in combined quantum mechanical and molecular mechanical (QM/MM) calculations of periodic systems. The QM/MM-Ewald method is a linear-scaling electrostatic method that utilizes the particle mesh Ewald algorithm for calculation of point charge interactions of molecular mechanical atoms and a real-space multipolar expansion for the quantum mechanical electrostatic terms plus a pairwise periodic correction factor for the QM and QM/MM interactions that does not need to be re-evaluated during the self-consistent field procedure. The method is tested in a series of molecular dynamics simulations of the ion−ion association of ammonium chloride and ammonium metaphosphate and the dissociative phosphoryl transfer of methyl phosphate and acetyl phosphate. Results from periodic boundary molecular dynamics (PBMD) simulations employing the QM/MM-Ewald method are compared with corresponding PBMD simulations using electrostatic cutoffs and with results from nonperiodic stochastic boundary molecular dynamics (SBMD) simulations, with cutoffs and with full electrostatics (no cutoff). The present method allows extension of linear-scaling Ewald methods to molecular simulations of enzyme and ribozyme reactions that use combined QM/MM potentials.

1. Introduction

The profound effects of solvation on chemical reactions have been recognized for over a century and continue to attract intensive experimental and theoretical research effort.[1] For reactions catalyzed by enzymes or ribozymes, the environment is even more complicated. The challenge in theoretical studies of the mechanism and reactivity of chemical processes is to move accurate quantum electronic structure calculations from the gas phase into the condensed phase realm. However, for large systems such as proteins and nucleic acids, the complexity and system size preclude the use of even the most efficient linear-scaling electronic structure methods[2] to simulate the reaction dynamics explicitly. This is further exacerbated by the need for an adequate treatment of long-range electrostatic interactions in polar solvents and in the presence of mobile counterions. Fortunately, it is often the case that the vast majority of the system does not require a high-level and computationally intensive quantum mechanical model. This situation is ideally suited for application of a combined quantum mechanical and molecular mechanical (QM/MM) approach, in which the solute is treated quantum mechanically and the environment by classical force fields.[3−6]

Electrostatic interactions are generally perceived to be the dominant forces that stabilize transition states in biochemical reactions[7−10] and provide essential stability in long-time dynamic simulations of proteins and nucleic acids.[11−16] The development of efficient linear-scaling electrostatic methods[14,17,18] (methods for which the computational effort scales linearly with system size − or nearly so, see below) have greatly improved the reliability of molecular dynamics simulations of large biological systems. In combined QM/

---

* Corresponding author e-mail: york@chem.umn.edu. Corresponding address: Department of Chemistry, University of Minnesota, Minneapolis, MN 55455-0431.

Efficient Linear-Scaling Ewald Method

J. Chem. Theory Comput., Vol. 1, No. 1, 2005 **3**

MM potential models, the electrostatic environment affects the quantum electronic polarization of the solute[19] that plays a significant role in the stabilization of macromolecules in solution[20] and the rate enhancement of some enzymes.[21] Consequently, it is critical to compute long-range electrostatic interactions accurately in QM/MM simulations of biochemical reactions. Nonetheless, due to the lack of availability of algorithms that extend linear-scaling electrostatic methods to combined QM/MM potentials, a large percentage of QM/MM applications routinely employ electrostatic cutoffs.[10]

The present paper presents a linear-scaling Ewald method for efficient calculation of long-range electrostatic interactions in combined QM/MM simulations using semiempirical quantum models. The method can be easily extended to QM/MM ab initio molecular orbital and density functional theory. In the following, the essential theory and computational details are first outlined. Next, results obtained using the QM/MM-Ewald method are compared with those from other simulations by computing interionic potentials of mean force for ion association and for phosphoryl transfer reactions. Specifically, periodic boundary molecular dynamics (PBMD) simulations calculated with the QM/MM-Ewald method are compared with corresponding PBMD simulations using electrostatic cutoffs and with results from nonperiodic stochastic boundary molecular dynamics (SBMD) simulations, with cutoffs and with full electrostatics (no cutoff). Finally, the paper concludes with a summary of the key results and identifies directions of future research.

2. Theory

2.1. Electrostatic Energy of a Periodic System of Point Charges.

Consider a periodic system of N point charges $\{q_i\}$ located at position $\{\mathbf{r}_i\}$, $i = 1, \cdots, N$, in a periodic unit cell U characterized by the set of real-space lattice vectors $\{\mathbf{a}_k\}$, $k = 1, 2, 3$. The classical electrostatic energy of this system, excluding the infinite self-energy of the point charges, is given by

$$E_{\text{elec}} = \frac{1}{2} \sum_{i}^{N} \sum_{j}^{N} {\sum_{\mathbf{n}}^{\infty}}' \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{n}|} \quad (1)$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, and the summation over $\mathbf{n}$ is over all integer translations of the real space lattice vectors $\mathbf{n} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$ for integers n_k ($k = 1, 2, 3$), and the prime symbol indicates that the terms where $|\mathbf{r}_{ij} + \mathbf{n}| = 0$ are neglected. The summation in eq 1 is not convergent unless the total charge of the system sums to zero (i.e., the monopole moment of the unit cell vanishes). If the unit cell has vanishing monopole *and* dipole moments, the sum converges absolutely; however, if the unit cell has a net dipole moment, the sum is only conditionally convergent and has different converged values depending on the order and limiting manner whereby the sum is affected.[22] In any case, the expression in eq 1, under the conditions where it does converge, does so very slowly, and is not a practical means of computing electrostatic energies for periodic systems.

The Ewald summation convention[23,24] uses an elegant mechanism of transforming the slowly convergent sum in

eq 1 into two rapidly convergent sums over real-space and reciprocal space lattice vectors

$$E_{\text{elec}}(q^N, \mathbf{r}^N, U) = E_{\text{real}}(q^N, \mathbf{r}^N, U; \kappa) + E_{\text{recip}}(q^N, \mathbf{r}^N, U; \kappa) + E_{\text{surf}}(q^N, \mathbf{r}^N, U; P, \epsilon) \quad (2)$$

where

$$E_{\text{real}}(q^N, \mathbf{r}^N, U; \kappa) = \frac{1}{2} \sum_{i}^{N} \sum_{j}^{N} q_i q_j \left({\sum_{\mathbf{n}}^{\infty}}' \frac{\text{erfc}(\kappa|\mathbf{r}_{ij} + \mathbf{n}|)}{|\mathbf{r}_{ij} + \mathbf{n}|} - \frac{2\kappa}{\sqrt{\pi}} \delta_{ij} \right) \quad (3)$$

$$E_{\text{recip}}(q^N, \mathbf{r}^N, U; \kappa) = \frac{1}{2} \cdot \frac{4\pi}{V} \sum_{|\mathbf{k}| \neq 0}^{\infty} \frac{\exp(-k^2/4\kappa^2)}{k^2} |S(\mathbf{k})|^2 \quad (4)$$

where $\text{erfc}(x)$ is the complementary error function, defined as $\text{erfc}(x) = 1 - \text{erf}(x)$, and $\text{erf}(x)$ is the error function.[25] The summation in eq 4 is over vectors $\mathbf{k} = 2\pi\mathbf{m}$, and $\mathbf{m}$ sums over all integer translations of the reciprocal lattice $\mathbf{m} = m_1\mathbf{a}_1^* + m_2\mathbf{a}_2^* + m_3\mathbf{a}_3^*$ for integers m_k ($k = 1, 2, 3$), where the set of reciprocal lattice vectors $\{\mathbf{a}^*_i\}$ are related to the real-space lattice vectors $\{\mathbf{a}_i\}$ by $\mathbf{a}_i^* \cdot \mathbf{a}_j = \delta_{ij}$. In eq 4, V is the volume of the unit cell U ($V = |\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3|$), and $S(\mathbf{k})$ is the structure factor[26,27] and is given by

$$S(\mathbf{k}) = \sum_{j}^{N} q_j \exp(i\mathbf{k} \cdot \mathbf{r}_j) \quad (5)$$

The two summations contain a parameter κ that adjusts the relative rates of convergence. The total energy is independent of the κ parameter, so long as the real-space and reciprocal space sums are both sufficiently converged. In practice, these sums are truncated at some point so as to fall below a fixed tolerance level in accuracy ϵ_{tol}. If the parameter κ is chosen such that only the $|\mathbf{n}| = 0$ term is required in eq 3 to obtain the desired level of accuracy (i.e., the *minimum image convention*[24,28] can be used to perform the summations over particles in the unit cell), then the number of reciprocal-space lattice vectors $\mathbf{k}$ required to obtain the same level of accuracy becomes constant with respect to scaling of the unit cell, and an order N^2 algorithm results.[26,29] If the parameter κ is optimized for scaling efficiency, then an order $N^{3/2}$ algorithm can be obtained.[30] However, to extend the method to very large systems, a so-called "linear-scaling" algorithm is required whereby the scaling is *better than* order N^λ, $\forall \lambda > 1$. Such algorithms have been developed previously,[18] perhaps the most commonly employed algorithm in molecular dynamics simulations is the particle mesh Ewald method[27,31] that has recently been extended to higher order multipole moments.[32]

Before proceeding further, it is worthwhile to briefly clarify further eq 4. It was mentioned previously that the original expression of eq 1 was subject to several convergence restrictions. These restrictions manifest themselves through the $|\mathbf{k}| = 0$ term, that in eq 4 has been neglected. Clearly care must be taken with the $|\mathbf{k}| = 0$ term since the sum involves a $1/k^2$ factor, that must be resolved via a limiting procedure involving the ratio $|S(\mathbf{k})|/|k|$. From eq 5

it is clear that $|S(0)| \neq 0$ for a non-neutral system, consistent with the statement earlier that eq 1 is nonconvergent under this condition. If the system is neutral, but the unit cell has a net dipole moment ($\mathbf{D} = \sum_j q_j \mathbf{r}_j$), then the $|\mathbf{k}| = 0$ term gives rise to the surface term, $E_{\text{surf}}(q^N, \mathbf{r}^N, U; P, \epsilon)$ in eq 2, that depends quadratically on the dipole moment $\mathbf{D}$

$$E_{\text{surf}}(q^N, \mathbf{r}^N, U; P, \epsilon) = \frac{1}{2}\alpha(P, \epsilon)|\sum_j q_j \mathbf{r}_j|^2$$

$$= \frac{1}{2}\alpha(P, \epsilon) \cdot |\mathbf{D}|^2 \qquad (6)$$

where the proportionality constant $\alpha(P, \epsilon)$ depends on the macroscopic shape of the crystal, P, and the dielectric constant, ϵ, of the surrounding medium.[22] The physical interpretation of the surface term is that of an energy associated with a dipole layer on the surface of the crystal, embedded in a polarizable dielectric medium. In the limit that the surrounding dielectric constant becomes infinite (i.e., is a conductor), the energy of the surface dipole layer vanishes, i.e.,

$$\lim_{\epsilon \to \infty} \alpha(P, \epsilon) = 0 \qquad (7)$$

In the literature, this is often referred to as employing "tin-foil" boundary conditions.[33] A number of studies have investigated whether the inclusion of the surface term is a physically reasonable model for macroscopic systems[22,34] (since in a real crystals instantaneous microscopic fluctuations of the unit cell dipole moment are not propagated synchronously to the macroscopic limit). For most simulations, this terms has little overall effect and is generally neglected. Consequently, this term will not be further discussed, and henceforth the assumption will be made that the sum in reciprocal space can be made neglecting the $|\mathbf{k}| = 0$ term (and hence the surface energy correction).

It is useful to note that the Ewald sum energy, for a neutral system, can be written in terms of a pair potential $\psi_E(\mathbf{r}_{ij})$ as[11,31]

$$E_{\text{elec}} = \frac{1}{2} \sum_i^N \sum_j^N q_i q_j \psi_E(\mathbf{r}_{ij}) \qquad (8)$$

where the Ewald pair potential, $\psi_E(\mathbf{r}_{ij})$, (assuming the Ewald parameter κ is chosen such that the summation over real-space lattice vectors includes only the $|\mathbf{n}| = 0$ term) is given by

$$\psi_E(\mathbf{r}_{ij}) = \left(\frac{(1 - \delta_{ij})\text{erfc}(\kappa|\mathbf{r}_{ij}|)}{|\mathbf{r}_{ij}|} - \frac{2\kappa}{\sqrt{\pi}}\delta_{ij} + \frac{4\pi}{V} \sum_{|\mathbf{k}| \neq 0} \frac{\exp(-k^2/4\kappa^2)}{k^2} \cos(\mathbf{k} \cdot \mathbf{r}_{ij}) \right) \qquad (9)$$

In fact, efficient Ewald sum algorithms have been designed that precompute the Ewald pair potential on a 3-dimensional grid and use multidimensional interpolation procedures to allow rapid evaluation in molecular dynamics simulations.[18,28,35] This procedure, although fast, still scales as order

N^2 and hence becomes limiting for large systems. However, as will be seen shortly, for hybrid QM/MM calculations where the QM part of the system is small, and update of the QM contribution to the Ewald energy is required at each step of an SCF procedure, the use of a correction to the Ewald pair potential becomes computationally efficient.

2.2. Electrostatic Energy of a Periodic System with a Smooth Charge Density. The focus of the present paper is to develop a linear-scaling method for efficient calculation of electrostatic interactions specifically for hybrid QM/MM calculations. The case of QM/MM calculations is somewhat specialized in that the quantum mechanical region is typically fairly small in relation to the much larger surrounding molecular mechanical environment. It would be considerably costly to Fourier transform directly the localized QM density that would require many reciprocal space lattice vectors (or alternatively, a very fine fast Fourier transform grid) in a typically very large QM/MM unit cell. On the other hand, for semiempirical QM models, there are very efficient methods for solution of the Poisson equation for the QM charge distribution in real space.[36,37] As will be discussed in more detail below, the object of the present work is to develop a method that takes advantage of the specialized features of QM/MM calculations and capitalizes simultaneously on the most efficient methods for calculating electrostatics of point charge and smooth density distributions.

To facilitate development of the method, the following general notation is introduced for the electrostatic interaction energy between two generalized charge distributions, Q_A and Q_B, under real-space nonperiodic boundary conditions (RS) and periodic boundary conditions (PB) as

$$E^X[A, B] = \frac{2 - \delta_{A,B}}{2} \int \int Q_A(\mathbf{r}) G^X(\mathbf{r}, \mathbf{r}') Q_B(\mathbf{r}') d^3r d^3r' \qquad (10)$$

where $G^X(\mathbf{r}, \mathbf{r}')$ is the generalized Green's function for the Poisson equation that is a solution of

$$\nabla_{\mathbf{r}}^2 G(\mathbf{r}, \mathbf{r}') = -4\pi\delta(\mathbf{r} - \mathbf{r}') \qquad (11)$$

and the superscript "X" in eq 10 specifies the boundary conditions, which can be either "RS" or "PB" for real-space or periodic boundary conditions, respectively. The term $\delta_{A,B}$ is equal to 1 when the charge distributions are the same, and zero when the charge distributions are different. In the case that Q_A and Q_B are the same charge distribution and that charge distribution contains point charges, it is further assumed that the infinite self-energy of the point charges are neglected. In short, $E^X[A, B]$ in eq 10 represents the normal classical electrostatic energy of the charge distribution Q_A interacting with the charge distribution Q_B, including the possibility that Q_A and Q_B are identical. Note also that, by this definition, $E^X[A, B] = E^X[B, A]$ and $E^X[A + B, A + B] = E^X[A, A] + E^X[A, B] + E^X[B, B]$.

For a QM/MM calculation, the charge distribution is partitioned into a QM charge distribution that consists of the quantum mechanical electron density and nuclear core charges and an MM charge distribution that consists of the partial atomic charges of the MM environment. The total

Efficient Linear-Scaling Ewald Method

J. Chem. Theory Comput., Vol. 1, No. 1, 2005 **5**

energy of the system, under periodic boundary conditions, is thus given by

$$E^{PB}[\rho + \mathbf{q}, \rho + \mathbf{q}] = E^{PB}[\rho,\rho] + E^{PB}[\rho,\mathbf{q}] + E^{PB}[\mathbf{q},\mathbf{q}] \quad (12)$$

where ρ represents the distribution of electron density ρ of the QM atoms (plus the core nuclear charges) and $\mathbf{q}$ represents the distribution of classical MM point charges. Recall that the number of atoms associated with the QM charge distribution (N_{QM}) is typically much smaller than the number of atoms associated with the MM charge distribution (N_{MM}) and that the latter distribution spans a much greater spatial extent. The main problem to overcome involves the calculation of the electrostatics in the $E^{PB}[\rho,\rho]$ and $E^{PB}[\rho,\mathbf{q}]$ terms, since, for semiempirical methods, these terms involve a smooth charge density with high atomic multipolar character. However, the same electrostatic interactions for this term are straightforward to calculate in real-space and, in fact, are part of the computational machinery of any stand-alone or integrated semiempirical quantum method. This observation motivates rewriting the expression of eq 12 as

$$E^{PB}[\rho + \mathbf{q}, \rho + \mathbf{q}] = (E^{PB}[\rho,\rho] - E^{RS}[\rho,\rho]) + E^{RS}[\rho,\rho] + $$
$$(E^{PB}[\rho,\mathbf{q}] - E^{RS}[\rho,\mathbf{q}]) + E^{RS}[\rho,\mathbf{q}] + E^{PB}[\mathbf{q},\mathbf{q}] \quad (13)$$

Consider now an approximate quantum mechanical charge distribution, $\mathbf{Q}$, that is modeled as a set of auxiliary point charges such that the electrostatic potential closely represents that of the full QM charge distribution at distances on the order of the distance between crystal images. In the present work, simple Mulliken charges[38] are used for this purpose. However, alternate charge partitioning,[39−41] charge mapping,[42−45] or charge fitting[46,47] procedures could also be used as well. The Mulliken charge is particularly convenient to incorporate into the Fock operator, as is discussed below, owing to the simple linear relation with the single-particle density matrix. At short range, the potential due to these charges will deviate significantly from the exact quantum mechanical potential, but at distances on the order of a full unit cell translation away, the differences are very small. This motivates introduction of the following approximation for eq 13 as

$$E^{PB}[\rho + \mathbf{q}, \rho + \mathbf{q}] \approx (E^{PB}[\mathbf{Q},\mathbf{Q}] - E^{RS}[\mathbf{Q},\mathbf{Q}]) + E^{RS}[\rho,\rho] + $$
$$(E^{PB}[\mathbf{Q},\mathbf{q}] - E^{RS}[\mathbf{Q},\mathbf{q}]) + E^{RS}[\rho,\mathbf{q}] + E^{PB}[\mathbf{q},\mathbf{q}] \quad (14)$$

The above equation for the Ewald energy is useful for practical implementation into semiempirical QM/MM methods. It is clear, for example, that eq 14 requires evaluation of the periodic boundary energy only for the point charge distributions, $\mathbf{Q}$ and $\mathbf{q}$, whereas the more complicated exact quantum mechanical charge distribution, ρ, is required to be evaluated in real space.

2.3. Combined QM/MM Potential in Real Space. Combined QM/MM potential methods have been reviewed extensively elsewhere[5,6,48] and are only briefly outlined in this subsection. The effective Hamiltonian for the combined QM/MM potentials treated in the present work take the form

$$\hat{H}_{eff} = \hat{H}^0_{QM} + \hat{H}^{el}_{QM/MM} + \hat{H}^{vdW}_{QM/MM} + \hat{H}_{MM} \quad (15)$$

where $\hat{H}^0_{QM}$ is the Hamiltonian for the QM charge distribution represented as nuclei and electrons within the Born−Oppenheimer approximation and $\hat{H}_{MM}$ is the molecular mechanical potential of MM atoms. Two coupling terms, $\hat{H}^{el}_{QM/MM}$ and $\hat{H}^{vdW}_{QM/MM}$, represent interactions between QM and MM sites: $\hat{H}^{el}_{QM/MM}$ accounts for the electrostatic interactions of electrons and nuclei on QM atoms with point charges on MM sites, and $\hat{H}^{vdW}_{QM/MM}$ represents the short-range Pauli exchange repulsion and the long-range dispersion interactions and is modeled by a Lennard-Jones form. In real-space calculation, which is modified below to include long-range electrostatic interactions, the electrostatic interaction Hamiltonian, $\hat{H}^{el}_{QM/MM}$, given in eq 15 is written, in atomic units, as an exact interaction Hamiltonian of QM nuclei and electrons with MM atoms represented by partial point charges

$$\hat{H}^{el-RS}_{QM/MM} = -\sum_i^{N_{MM}} \sum_a^{N_e} \frac{q_i}{r_{ia}} + \sum_i^{N_{MM}} \sum_\alpha^{N_{QM}} \frac{q_i Z_\alpha}{R_{i\alpha}} \quad (16)$$

where q_i and Z_α are charges on MM and QM nuclei, N_e is the total number of electrons in QM region, and r_{ia} and $R_{i\alpha}$ are the distances of the quantum electrons and nuclei from the classical charge sites, respectively.

Then, the real-space potential energy in the combined QM/MM potential is computed using eq 17

$$E^{RS}[\rho] = \langle \Phi | \hat{H}_{eff} | \Phi \rangle = $$
$$E_{QM}[\rho] + E^{el-RS}_{QM/MM}[\rho,\mathbf{q}] + E^{vdW-RS}_{QM/MM} + E_{MM} \quad (17)$$

where Φ is the wave function of the solute in the field of MM environment, in which $E_{QM} + E^{el-RS}_{QM/MM}$ is determined through Hartree−Fock self-consistent-field (SCF) MO calculation by solving the Roothaan-Hall equation[49]

$$\mathbf{F}^{RS}\mathbf{C}^{RS} = \mathbf{S}\mathbf{C}^{RS}\mathbf{E}^{RS} \quad (18)$$

where $\mathbf{F}^{RS}$, $\mathbf{C}^{RS}$, and $\mathbf{S}$ denote the Fock, eigenvectors, and overlap matrices in real space, respectively, and $\mathbf{E}^{RS}$ is the diagonal matrix of orbital energies for molecular orbitals. The combined QM/MM potential constructed in this way is expected to be valid in the range that the QM model is adequately large so as to capture the essential chemical reaction process, and the MM model provides a sufficiently accurate representation of the electrostatic environment.

2.4. Ewald Modifications to the Effective Hamiltonian (Fock) Matrix Elements. The elements of the effective Hamiltonian (Fock) matrix, $F_{\mu\nu}$, in a periodic boundary system are defined as

$$F_{\mu\nu} = \frac{\delta E[\rho]}{\delta \rho_{\mu\nu}} \quad (19)$$

where $E[\rho]$ is the total energy that depends on the single-particle density matrix, ρ, with elements $\rho_{\mu\nu}$ and is related to the terms in eq 14 by

$$E[\rho] = (E^{PB}[\mathbf{Q},\mathbf{Q}] - E^{RS}[\mathbf{Q},\mathbf{Q}]) + E^{RS}[\rho,\rho] + $$
$$(E^{PB}[\mathbf{Q},\mathbf{q}] - E^{RS}[\mathbf{Q},\mathbf{q}]) + E^{RS}[\rho,\mathbf{q}] \quad (20)$$

This energy can be decomposed into QM and QM/MM

components, each of which consists of a real-space term plus a *periodic boundary correction* (PBC) as

$$E[\rho] = E_{QM}^{RS}[\rho] + \Delta E_{QM}^{PBC}[\mathbf{Q}] + E_{QM/MM}^{RS}[\rho] + \Delta E_{QM/MM}^{PBC}[\mathbf{Q}]$$
$$= E^{RS}[\rho] + \Delta E^{PBC}[\mathbf{Q}] \qquad (21)$$

where

$$E_{QM}^{RS}[\rho] = E^{RS}[\rho,\rho] \qquad (22)$$

$$\Delta E_{QM}^{PBC}[\mathbf{Q}] = (E^{PB}[\mathbf{Q},\mathbf{Q}] - E^{RS}[\mathbf{Q},\mathbf{Q}]) \qquad (23)$$

$$E_{QM/MM}^{RS}[\rho] = E^{RS}[\rho,\mathbf{q}] \qquad (24)$$

$$\Delta E_{QM/MM}^{PBC}[\mathbf{Q}] = (E^{PB}[\mathbf{Q},\mathbf{q}] - E^{RS}[\mathbf{Q},\mathbf{q}]) \qquad (25)$$

$$E^{RS}[\rho] = E_{QM}^{RS}[\rho] + E_{QM/MM}^{RS}[\rho] \qquad (26)$$

$$\Delta E^{PBC}[\mathbf{Q}] = \Delta E_{QM}^{PBC}[\mathbf{Q}] + \Delta E_{QM/MM}^{PBC}[\mathbf{Q}] \qquad (27)$$

Note that although the energy $E[\rho]$ depends on the density matrix ρ in a fairly complicated way, the periodic boundary correction terms (eqs 23 and 25) depend on the density matrix only through the atomic charge vector $\mathbf{Q}$. The Fock matrix can similarly be decomposed into a real-space term plus a periodic boundary correction as

$$F_{\mu\nu}^{PB} = F_{\mu\nu}^{RS} + \Delta F_{\mu\nu}^{PBC} \qquad (28)$$

where

$$F_{\mu\nu}^{RS} = \frac{\delta}{\delta\rho_{\mu\nu}}\{E_{QM}^{RS}[\rho] + E_{QM/MM}^{RS}[\rho]\} = \frac{\delta E^{RS}[\rho]}{\delta\rho_{\mu\nu}} \qquad (29)$$

$$\Delta F_{\mu\nu}^{PBC} = \frac{\delta}{\delta\rho_{\mu\nu}}\{\Delta E_{QM}^{PBC}[\mathbf{Q}] + \Delta E_{QM/MM}^{PBC}[\mathbf{Q}]\} = \frac{\delta\Delta E^{PBC}[\mathbf{Q}]}{\delta\rho_{\mu\nu}} \qquad (30)$$

For the purposes of the present work, only the *correction* to the Fock matrix ($\Delta F_{\mu\nu}^{PBC}$) that arises from introduction of the periodicity is described, because it is presumed that the machinery for construction of the complete Fock matrix in real space ($F_{\mu\nu}^{RS}$) is already available.[36,50] For the purposes of implementation, the present work formulates extension to periodic systems as an additional term that can be included by an auxiliary computer subroutine or module.

The task that remains is to write the periodic boundary correction to the energy in eq 27 in terms of the set of Mulliken charges[38] $\{Q_\alpha\}$, defined in NDDO-based semiempirical methods as

$$Q_\alpha = Z_\alpha - \sum_{\mu\in\alpha}\rho_{\mu\mu} \qquad (31)$$

where α is an atom index, in a manner that is efficient to calculate and update during the self-consistent field (SCF) procedure. As mentioned previously, other charge parti-

tioning,[39−41] charge mapping,[42−45] or charge fitting[46,47] methods may also be employed so long as a rigorous mapping to the single-particle density matrix can be affected such that the Fock matrix may be modified accordingly. From the periodic correction to the energy ($\Delta E^{PBC}[\mathbf{Q}]$ of eq 27), the chain relation is used to obtain the periodic correction to the Fock matrix elements as

$$\frac{\delta\Delta E^{PBC}[\mathbf{Q}]}{\delta\rho_{\mu\nu}} = \sum_{\alpha\in\mu,\nu}\frac{\delta Q_\alpha}{\delta\rho_{\mu\nu}}\cdot\frac{\partial\Delta E^{PBC}[\mathbf{Q}]}{\partial Q_\alpha} =$$
$$-\sum_{\alpha\in\mu,\nu}\delta_{\mu\nu}\cdot\frac{\partial\Delta E^{PBC}[\mathbf{Q}]}{\partial Q_\alpha} \qquad (32)$$

To maximize efficiency of the method, one must bear in mind that the number of MM atoms, N_{MM}, usually greatly exceeds the number of quantum atoms, N_{QM}. Moreover, at each molecular dynamics integration step, the calculation of the QM/MM total energy and gradient requires an SCF procedure to be performed, and consequently, the periodic potential due to the QM charge distribution must be updated (recalculated) at each SCF iteration. Note that the atomic positions of all the atoms in the system remain fixed during the SCF procedure.

Consider the first periodic energy correction term, $\Delta E_{QM}^{PBC}[\mathbf{Q}]$, of eq 23. If the convention is used that the Ewald parameter κ is chosen such that the summation over real-space lattice vectors includes only the $|\mathbf{n}| = 0$ term, the first periodic energy correction term can be written concisely as

$$\Delta E_{QM}^{PBC}[\mathbf{Q}] = \frac{1}{2}\sum_{\alpha}^{N_{QM}}\sum_{\beta}^{N_{QM}}Q_\alpha Q_\beta\Delta\psi_E(\mathbf{R}_{\alpha\beta}) \qquad (33)$$

where $\Delta\psi_E(\mathbf{R}_{\alpha\beta})$ is the periodic correction to the Ewald pair potential (eq 9), and with the choice of κ above, is given by

$$\Delta\psi_E(\mathbf{R}_{\alpha\beta}) =$$
$$\left(\frac{4\pi}{V}\sum_{|\mathbf{k}|\neq 0}\frac{\exp(-k^2/4\kappa^2)}{k^2}\cos(\mathbf{k}\cdot\mathbf{R}_{\alpha\beta}) - \frac{\operatorname{erf}(\kappa|\mathbf{R}_{\alpha\beta}|)}{|\mathbf{R}_{\alpha\beta}|}\right) \qquad (34)$$

Note that in the derivation of eq 34 the limiting relation was used

$$\lim_{r\to 0}\frac{\operatorname{erf}(\kappa r)}{r} = \frac{\kappa}{\sqrt{\pi}} \qquad (35)$$

For the QM periodic correction term (eq 33), the corresponding correction to the Fock matrix must be recalculated at each step of the SCF procedure. Since the number of quantum atoms is small, the correction to the Ewald pair potential (eq 34) needed for the Fock matrix can easily be calculated once as an $N_{QM} \times N_{QM}$ matrix and stored and hence not be recalculated during the SCF. This makes calculation of the periodic correction to the Fock matrix consist of a simple matrix multiplication of the Ewald pair potential correction with the Mulliken charge vector, only the latter of which changes at each iteration.

Efficient Linear-Scaling Ewald Method

J. Chem. Theory Comput., Vol. 1, No. 1, 2005 **7**

Consider now the second periodic energy correction term, $\Delta E_{QM/MM}^{PBC}[\mathbf{Q}]$, of eq 30 (see also eq 25) that can be written as

$$\Delta E_{QM/MM}^{PBC}[\mathbf{Q}] = \sum_{\alpha}^{N_{QM}} Q_{\alpha} \sum_{j}^{N_{MM}} q_j \Delta \psi_E(\mathbf{R}_{\alpha} - \mathbf{r}_j) \qquad (36)$$

In this case, the corresponding correction to the Fock matrix does *not* need to be updated during the SCF procedure, since the MM charge distribution is not changing. Consequently, the periodic correction to the static potential of the MM charges at the QM charge positions can be calculated once as a $N_{QM} \times 1$ vector and simply added to the 1-electron terms of the Fock matrix (sometimes referred to as the *core* Hamiltonian matrix). The above procedure leads to an efficient method for calculation of long-range electrostatic interactions in combined QM/MM calculations.

3. Computational Details

The combined QM/MM-Ewald sum method has been implemented into a modified version of CHARMM[51] (version c30a1) interfaced with the MNDO97 program[52] and MOPAC.[53] To test the method introduced in the present work, simulations of ion association processes and dissociative mechanisms of phosphoryl transfer were performed. The interionic potential of mean force was calculated for each simulation and compared with calculated values from nonperiodic (full electrostatic or cutoff) and periodic (cutoff) simulations. In particular, the following systems were examined: the ionic association of (1) ammonium chloride and (2) ammonium metaphosphate and the dissociative phosphoryl transfer mechanism of (3) methyl phosphate and (4) acetyl phosphate. The semiempirical AM1 model[50] was used for the ammonium chloride system, and MNDO/d[54] for the phosphorus-containing systems for which d-orbitals have been shown to be important.[55] The solutes were treated fully quantum mechanically at the semiempirical level and were solvated in a 40.0 Å cubic box of TIP3P water molecules,[56] resulting in a total of 2042 waters for the ammonium chloride system, 2040 waters for the ammonium metaphosphate system, 2038 waters for the methyl phosphate system, and 2035 waters for the acetyl phosphate system. Internal water geometries were constrained using the SHAKE algorithm in all simulations.[57]

A spherical cutoff scheme was used to evaluate the van der Waals and the real-space electrostatic interactions in the QM/MM-Ewald method and in the non-Ewald approaches. In all cases, water molecules were included in the cutoff list if the geometrical center of water was less than the cutoff distance from any group center of the solute. The solute ion pair was divided into two groups, one for each formal ion. It should be emphasized that in evaluating QM/MM interaction energies, solute−solvent (or QM-MM) interactions were determined for the entire QM system, whenever a solvent molecule was within the cutoff distance from any solute (QM) group. In simulations using the QM/MM-Ewald sum technique, a 10.0 Å group-based cutoff was used. The nonbonded list and crystal images were updated every 25 steps during molecular dynamics simulations. For Ewald

summation, the Ewald κ value (see above) was chosen to be 0.340 Å^{-1}, and the smooth particle mesh Ewald (PME) method was employed for reciprocal space summations between MM sites with an approximate grid size of 0.8 Å ($50 \times 50 \times 50$ FFT grid)[27,31] and with net charge correction to the Ewald potential suggested by Bogusz et al.[33] All simulations were propagated using the leapfrog Verlet algorithm with 1 fs integration time step.[24] Periodic boundary conditions were used along with the isothermal−isobaric ensemble (NPT) at 1 atm and 298 K using extended system pressure algorithm of Andersen[58] with effective mass of 500.0 amu and Hoover thermostat[59] with effective mass of 1000.0 kcal/mol-ps^2, respectively. In the QM/MM simulations under PB without Ewald summation, electrostatic interactions were determined using a spherical cutoff scheme based on group separation with switching between 10.5 and 11.5 Å.

The potential of mean force (PMF) profiles have been determined using umbrella sampling,[60] in which PMF is represented as a function of internuclear distance defined as the N···Cl distance (R_{N-Cl}) in ammonium chloride, the N···P distance (R_{N-P}) in ammonium metaphosphate, and the O···P distance (R_{O-P}) in methyl phosphate and acetyl phosphate. After initial 200 ps of equilibration, 25 separate umbrella sampling windows (28 windows in the simulations of dissociative phosphoryl transfer reactions) were executed to span the internuclear separation up to 12.0 Å by applying a harmonic restraining potential centered at the center of the particular umbrella window. The spacing between neighboring windows was a function of the interionic separation distance: 2.0, 2.5, and 5.0 Å spacings were used for interionic separation distances R in the range of $R \leq 3.0$ Å, 3.0 Å $\leq R \leq 4.0$ Å, and $R \geq 4.0$ Å, respectively. The force constants used were chosen and adjusted based on the shape of PMF profile for each system tested to guarantee sufficient overlap of the probability distribution with neighboring windows (force constant values ranged between 80.0−150.0 kcal/mol-Å in the region of steep repulsive wall at small R, 40.0−80.0 kcal/mol-Å in the intermediate separation, and 10.0−40.0 kcal/mol-Å in the region of large interionic separation). Each umbrella sampling window was equilibrated for 35 ps followed by 50 ps of production with data collected every step. The weighted histogram analysis method (WHAM)[61] was employed to compute the potential of mean force as a function of internuclear separation.

For additional comparison of the ion association simulations, umbrella sampling simulations were also performed to compute interionic PMF profiles with stochastic boundary molecular dynamics (SBMD)[62,63] by using spherical water box for ammonium chloride with 1034 TIP3P water molecules and ammonium metaphosphate with 1031 TIP3P waters. Simulations were performed without cutoff as well as with a 11.5 Å nonbonded cutoff as in the periodic simulations. The radius of water sphere is 20.0 Å to keep the size of the simulation fairly close to that of the PB simulations.

Table 1. Relative Fluctuations in Potential Energy (σ_E), Volume (σ_V), Temperature (σ_T), and Average Total Force (F_{ave}^{tot}) from MD Simulations Using QM/MM-Ewald with PME for MM Electrostatics[a]

	σ_E	σ_V	σ_T	F_{ave}^{tot} [b]
$NH_4^+\cdots CL^-$	3.60×10^{-3}	5.52×10^{-3}	1.01×10^{-2}	2.55×10^{-2}
$NH_4^+\cdots PO_3^-$	3.49×10^{-3}	4.77×10^{-3}	0.99×10^{-2}	2.51×10^{-2}
$CH_3O^-\cdots PO_3^-$	3.75×10^{-3}	4.04×10^{-3}	1.05×10^{-2}	2.51×10^{-2}
$CH_3CO_2^-\cdots PO_3^-$	3.47×10^{-3}	5.42×10^{-3}	1.03×10^{-2}	2.54×10^{-2}

[a] $\sigma_E = rms_E/E_{ave}$, $\sigma_V = rms_V/V_{ave}$, and $\sigma_T = rms_T/T_{ave}$ where rms is the root mean square deviation of those properties. [b] The average total force goes to zero when regular Ewald sum method is used for MM electrostatics.

4. Results and Discussion

The focus of the current paper is to describe the development and implementation of an efficient QM/MM-Ewald method and provide benchmark simulation tests and to characterize the effects of treatment of electrostatic interactions on the QM/MM free energy profiles. Emphasis will not be placed on detailed analysis of the simulations and comparison with experiment that would first require a more quantitative assessment of the accuracy of the quantum and solvation models and QM/MM parameters. Development of new semiempirical quantum models for chemical reactions is an area of intense effort,[64−70] and consequently the extension of methods for efficient treatment of long-range electrostatic interactions in these calculations is of prime importance.

This section presents results of simulations of ion association and dissociative phosphoryl transfer using the QM/MM-Ewald method developed in the present work. The effects of periodicity and electrostatic cutoff are compared from PMF profiles of separate simulations. For the ion association tests (ammonium chloride and ammonium metaphosphate), PBMD and SBMD simulations are compared both with and without electrostatic cutoffs (the simulations without cutoffs are referred to as the QM/MM-Ewald and full-electrostatic SBMD for periodic and nonperiodic systems, respectively). For the phosphoryl transfer reactions (methyl phosphate and acetyl phosphate), PBMD simulations with and without electrostatic cutoff are compared (i.e., PBMD with cutoff and QM/MM-Ewald simulations). A summary of the observed fluctuations in potential energy, volume, temperature, and total force over a 10 ps interval for each of the QM/MM-Ewald simulations are shown in Table 1. The relative force errors for cutoff and alternative Ewald methods with respect to the QM/MM-Ewald method are compared in Table 2. A noteable feature is that inclusion of the Ewald term as a post-SCF correction still leads to considerable force error. The comparison underscores the importance of inclusion of long-range electrostatics in the simulation and updates of the periodic contribution to the quantum polarization during the SCF procedure since, although the post-SCF molecular mechanical Ewald correction of QM/MM calculation using Mulliken charge representation on quantum atoms improves the force error significantly, it still leads to considerable error.

4.1. Association of Oppositely Charged Ions. 4.1.1. Ammonium Chloride. Figure 1 compares the PMF profiles for ion association of ammonium chloride as a function of the N···Cl distance from simulations with and without electrostatic cutoffs under nonperiodic and periodic boundary conditions. The zeros of the PMFs were set to the minimum value for the ionic complex.

Table 2. Comparison of the Force Errors on the QM Atoms for Several Electrostatic Methods: Root Mean Square Error (RMSE), Mean Signed Error (MSE), Mean Unsigned Error (MUE) of Force, and Maximum Force Error (MAXE)[a]

	RMSE	MSE	MUE	MAXE
	Methyl Phosphate[b]			
10.0 Å cutoff	1.618	−0.010	1.221	4.612
11.5 Å cutoff	2.221	0.052	1.726	6.073
Post-SCF Ewald[c]	1.155	0.006	0.808	3.864
	Acetyl Phosphate[d]			
10.0 Å cutoff	2.836	−0.064	1.961	8.404
11.5 Å cutoff	1.830	−0.015	1.195	5.271
Post-SCF Ewald[c]	1.810	−0.009	1.138	6.642

[a] All units are kcal·mol^{-1}·Å^{-1}. [b] The forces are computed at the transition state of $R_{O-P} = 3.2$ Å. [c] The forces from Ewald potential are added to the forces computed from 10.0 Å cutoff, in which the Ewald potential has been computed from the Mulliken charges of the QM atoms. [d] The forces are computed at the transition state of $R_{O-P} = 2.9$ Å.
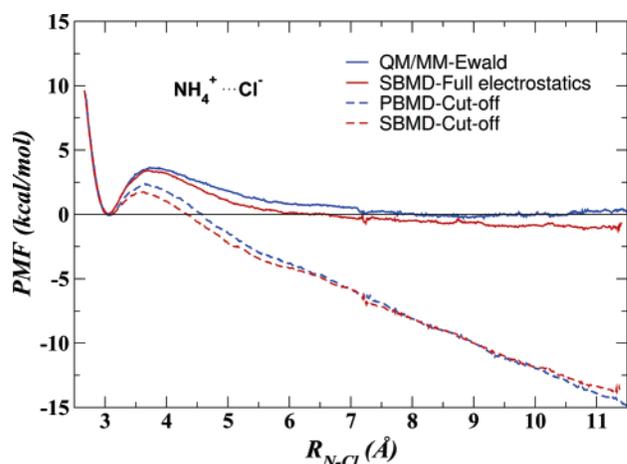
Figure 1. Comparison of potential of mean force (PMF) profiles for the ionic separation (R_{N-Cl}) of ammonium chloride ($NH_4^+\cdots Cl^-$) in water. Profiles were constructed from MD simulations with periodic boundary conditions using the combined QM/MM-Ewald sum (solid blue line) and PBMD method with 11.5 Å cutoff (dashed blue line) and with full-electrostatic SBMD (solid red line) and SMBD with 11.5 Å (dashed red line).

The electrostatic cutoff causes an artificial decrease in the PMF for oppositely charged ions at a large separation irrespective of whether the system is treated with spherical stochastic boundary (SBMD with cutoff) or with periodic boundary (PBMD with cutoff) conditions. Alternately, the full-electrostatic simulations (QM/MM-Ewald and full-

Efficient Linear-Scaling Ewald Method

J. Chem. Theory Comput., Vol. 1, No. 1, 2005 **9**

electrostatic SBMD) show the expected flattening of the PMF at large distances. These profiles are similar between the QM/MM-Ewald and the full-electrostatic SBMD simulations. The behavior of all the PMF profiles are similar up to an interion separation of around 3.0 Å, after which, the results of the cutoff simulations begin to diverge significantly from those of the QM/MM-Ewald and full-electrostatic SBMD simulations. The barrier for ionic dissociation is about 3.6 kcal/mol (QM/MM-Ewald), 3.4 kcal/mol (full-electrostatic SBMD), 2.3 kcal/mol (PBMD with cutoff), and 1.7 kcal/mol (SBMD with cutoff), respectively. The cutoff methods affect the energy barrier by over 1 kcal/mol at the transition state (a distance of less than 0.8 Å from the minimum) and have an even more profound effect at larger distances in the PMF profile.

At large separation, the PMF profiles of the simulations with long-range electrostatics become relatively flat after about 7.0 Å, indicating the ions are effectively shielded by the nonlocal solvent response. Alternately, the PMF profiles of the cutoff simulations show a steady linear drift from around 6.0 Å out past 10 Å. This linear drift of the PMF profiles for the cutoff simulations is due to an imbalance in the electrostatic interactions. The dipole moments of the waters solvating the individual oppositely charged ions are favorably aligned in the region between the ions and unfavorably aligned at opposite ends. The unfavorable interactions at the ends fall outside the cutoff first as the ions separate, while the interactions of the favorably aligned waters in the center are retained and result in the artificial drift in the PMF profiles. It is likely, therefore, that QM/MM simulations of biochemical reactions that involve association or dissociation of oppositely charged species, such as seen in many biochemical S_N1 reactions and photodissociation processes, may be subject to artificial overstabilization of the separated ionic species if electrostatic cutoffs are used.

4.1.2. Ammonium Metaphosphate. Figure 2 compares the PMF profiles for ion association of ammonium metaphosphate as a function of the N···P distance from simulations with and without electrostatic cutoffs under nonperiodic and periodic boundary conditions. The zeros of the PMFs were set to the limiting long-range value for the QM/MM-Ewald and full-electrostatic SBMD simulations and adjusted such that the short-ranged repulsive wall were coincident for the cutoff simulations. In these simulations, the MNDO/d Hamiltonian was employed since it has been demonstrated to provide a reliable description of biological phosphorus compounds.[54,55,70] Unlike the ammonium chloride PMF (Figure 1) that exhibits a stable free energy minimum for the ion−ion complex, the PMF profile for ammonium metaphosphate decreases monotonically. In the short range (N···P distances less than 4.0 Å), all of the PMF profiles are similar; however, after 4.0 Å, the PMF values for the cutoff simulations diverge from those of the QM/MM-Ewald and full-electrostatic SBMD simulations. The long ranged behavior of the PMF for the QM/MM-Ewald and full-electrostatic SBMD are quite similar, exhibiting a flat asymptotic limit after around 6.5 Å, indicating the oppositely charged ions are effectively screened. The PMF profiles for
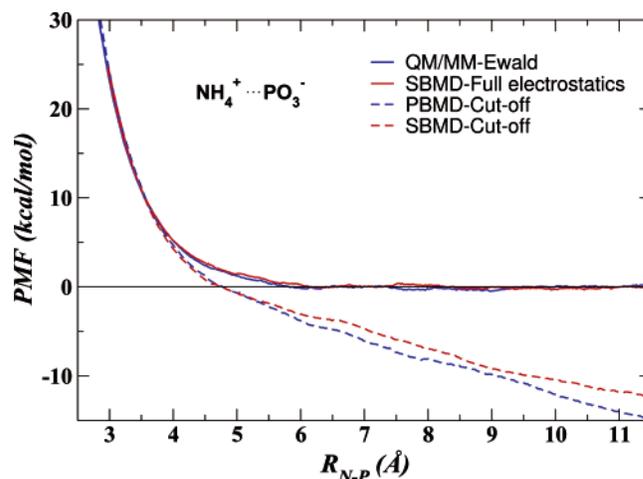

Figure 2. Comparison of potential of mean force (PMF) profiles for the ionic separation (R_{N-P}) of ammonium meta-phosphate (NH_4^+···PO_3^-) in water. Profiles were constructed from MD simulations with periodic boundary conditions using the combined QM/MM-Ewald sum (solid blue line) and PBMD simulations with 11.5 Å cutoff (dashed blue line) and with full-electrostatic SBMD (solid red line) and SMBD with 11.5 Å (dashed red line).

the PBMD and SBMD cutoff simulations show a linear drift after 6.0 Å, as in the ammonium chloride case (Figure 1). It is likely that the metaphosphate plane has a random orientation relative to the ammonium ion at large separation, but, as the ions approach each other, the plane is aligned perpendicular to the N···P vector. At this short separation, the partial positive nature of phosphorus atom cancels the favorable interactions between ammonium and oxygens in metaphosphate and results in no stable ion−ion complex. The results of these simulations echo those for the ammonium chloride system: QM/MM simulations of reactions that involve the dissociation of oppositely charged ions may lead to separated ionic species that are significantly overstabilized.

4.1.3. Effect of Cutoff on the Association of Oppositely Charged Ions. The effect of treatment of electrostatic interactions for the association of oppositely charged ions has been studied previously, although to our knowledge, not with the same QM/MM-Ewald model as presented in the present work. At long range, both ion association PMF profiles exhibit an artificial linear drift using either PBMD or SBMD with cutoff. The slope of the linear drift from 6 to 10.0 Å is similar between the PBMD and SBMD cutoff simulations with values of −2.25 and −1.66 kcal/mol-Å, respectively, for ammonium chloride and −2.01 and −1.86 kcal/mol-Å, respectively, for ammonium metaphosphate. For opposite-charged ionic systems, the results from current simulations indicate an unphysical roughly linear downward drift in the PMF profiles as the ions separate. Rozanska and Chipot also observed a similar artifact for the PMF profile of guanidinium-acetate association from molecular dynamics simulations using a smoothed electrostatic cutoff, whereas the simulations using the Ewald sum showed the expected solvent shielded behavior of solvent-separated ion pair, and the generalized reaction field correction also significantly
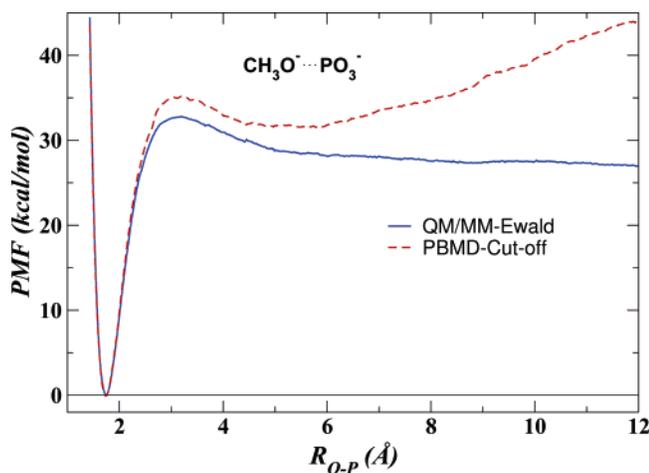
Figure 3. The computed PMF from PBMD simulations using combined QM/MM-Ewald sum potential (solid blue line) and with 11.5 Å cutoff (dashed red line) for the dissociation of methyl phosphate ($CH_3O^-\cdots PO_3^-$) in water.

improved the description relative to the cutoff method except toward the edge of the cutoff sphere.[71]

4.2. Dissociative Phosphoryl Transfer Mechanisms. The nonenzymatic and enzymatic chemical mechanism of phosphate hydrolysis reactions remains a topic of discussion and considerable debate.[72] Phosphate hydrolysis reactions are often discussed in terms of their associative or dissociative character[73] that can sometimes be distinguished kinetically. Kinetic measurements provide crucial data for these reactions, although they do not always provide a unique mechanistic interpretation.[74,75] The associative versus dissociative character is governed by many factors including the degree of esterification of the phosphate, the protonation state, the nature of the leaving group, and interactions with solvent, ions, and macromolecular environment.

To test the QM/MM-Ewald method, the dianionic dissociative phosphoryl transfer pathways for methyl phosphate and acetyl phosphate were examined using PBMD simulations both with electrostatic cutoff and with the QM/MM-Ewald method. Methyl phosphate is a commonly employed model for phosphoryl transfer reactions in kinases and phosphatases, and acetyl phosphate represents a model for a high-energy intermediate in the metabolism of many bacteria.[76] The first step of the dissociative pathway (Scheme 1) involves a dephosphorylation step characterized by the departure of a solvated metaphosphate (PO_3^-) group. The second step of the reaction involves the nucleophilic substitution to the metaphosphate by a nucleophile (usually a hydroxide ion or a water molecule in solution). This type of reaction is referred to as a $D_N + A_N$ type mechanism in IUPAC nomenclature.[77]

4.2.1. Methyl Phosphate. Figure 3 shows the computed PMF profile for methyl phosphate. The free energy of dissociation from current PMF profiles with QM/MM-Ewald method is 27.1 kcal/mol for methyl phosphate, and the activation free energy barrier is 32.8 kcal/mol. The effect of cutoff in the PBMD simulations is to raise the activation free energy barrier to 35.4 kcal/mol (an increase of 2.6 kcal/mol, or 8%). The effect of cutoff is even more pronounced

Scheme 1. Dissociative Phosphoryl Transfer Mechanism

Methyl phosphate

Acetyl phosphate

on the free energy of dissociation due to cutoff artifacts of the like-charged ions at fairly large separation (see below).

There have been several studies from experiment and theory for the dissociative reaction of methyl phosphate. The reaction free energy estimated from experiment by Guthrie is 37 ± 3 kcal/mol.[75,78] This indicates the reaction free energy calculated with the present work may be as much as 10 kcal/mol in error. Although it is not the purpose here to present free energy profiles with the greatest accuracy, it is worthwhile to point out the likely sources of error in order to assist in the development of improved QM/MM models. The main sources of error involve the semiempirical quantum model itself, the simplistic molecular mechanical model for water, and the QM/MM van der Waals interactions. The latter has a tremendous effect on the reaction free energies and barrier heights for processes that involve ion association/dissociation. For example, the heat of formation of methoxide (CH_3O^-) ion computed from MNDO/d Hamiltonian gives -39.7 kcal/mol in the gas phase,[36,54] while the experimentally determined value is -32.2 ± 1.1 kcal/mol.[79] The error in the methoxy ion from semiempirical MNDO/d model alone is 7.5 kcal/mol, which is close to the 10 kcal/mol difference between computed reaction free energy and experiment. A promising approach toward improvement of semiempirical quantum models is to develop reaction-specific parameters[70] that closely reproduce high-level quantum results.[80,81] Additionally, the model for water that was employed lacks explicit electronic polarizability, which is expected to be important for the stabilization of highly ionic systems such as those studied here. Finally, the optimization of the QM/MM van der Waals radii[82−84] to reproduce correct relative solvation free energies is critical. All of these areas will be addressed in future work.

4.2.2. Acetyl Phosphate. Figure 4 shows the computed PMF profiles for acetyl phosphate. The free energy of dissociation from current PMF profiles with QM/MM-Ewald method is 6.8 kcal/mol for acetyl phosphate, and the activation free energy barrier is 12.2 kcal/mol. The effect of cutoff in the PBMD simulations raised the activation free energy barrier to 13.8 kcal/mol (an increase of 1.6 kcal/mol or 13%). As for the dissociative phosphoryl transfer of methyl phosphate, the effect of cutoff has an even more profound effect on the free energy of dissociation (see below).
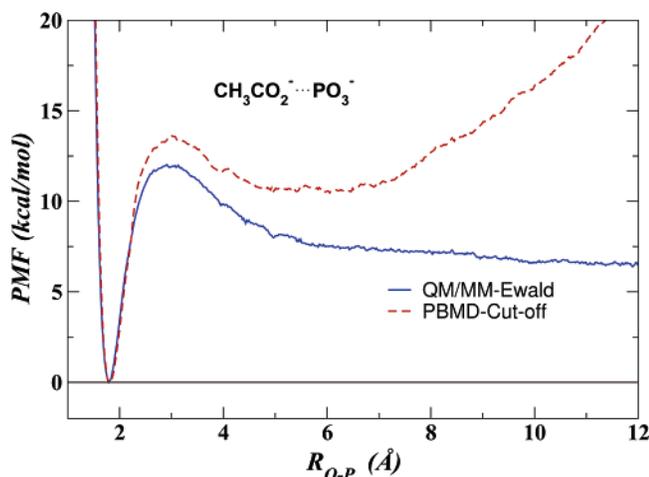
Efficient Linear-Scaling Ewald Method

J. Chem. Theory Comput., Vol. 1, No. 1, 2005 **11**

Figure 4. The computed PMF from PBMD simulations using combined QM/MM-Ewald sum potential (solid blue line) and with 11.5 Å cutoff (dashed red line) for the dissociation of acetyl phosphate ($CH_3CO_2^-\cdots PO_3^-$) in water.

The dissociative free energy activation barrier for acetyl phosphate is predicted to be 20.6 kcal/mol less than that for methyl phosphate. This is largely due to the increased stability of the acetate anion in solution, which has a much lower pK_a value (4.8) than methanol (15.5),[85] and is a considerably better leaving group. The reaction free energy for acetyl phosphate is similarly predicted to be lower than that of methyl phosphate by 20.3 kcal/mol. Thus, the lowering of the activation barrier can be explained by the added stabilization of the acetate anion relative to the methoxide anion, in accord with the Hammond postulate.[86,87]

4.2.3. Effect of Cutoff on the Dissociation of Like Charged Ions. The free energy of dissociation from current PMF profiles with Ewald sum is 27.1 kcal/mol for methyl phosphate and 6.8 kcal/mol for acetyl phosphate, respectively. The barrier height is 32.8 kcal/mol for methyl phosphate and 12.2 kcal/mol for acetyl phosphate from simulations with Ewald sum, while it is 35.4 kcal/mol and 13.8 kcal/mol with the cutoff method, respectively. The radial distribution function of water molecule around the solute at the transition state of both reactions has been checked, but no significant differences between QM/MM-Ewald and cutoff simulations were observed (data not shown). Thus, the difference in the barrier heights of dissociation reactions can most likely be attributed to long-range electrostatic effects that involve the ions and solvent.

The PMF profiles of the like-charged ionic systems of the present work are nonmonotonic and exhibit a broad minimum between 5 and 7 Å. Nonmonotonic PMF profiles arising from the use of electrostatic cutoffs have been previously observed by Bader and Chandler[88] in the dissociation of ferric and ferrous ion models in aqueous solution, whereas with the use of Ewald sums, correct monotonic PMF profiles were obtained. Furthermore, a comparison of the spatial distribution functions of like ion pairs have been investigated by Dang and Pettitt using molecular dynamics simulation and integral equation theory.[89] The results suggest the existence of a minimum for Cl^-–Cl^- pair at close distances and also a slight minimum at larger distance (6–7 Å) using a cutoff

method with switching.[89] Del Buono et al. studied Cl^-–Cl^- pair and ferric and ferrous ion pair by computing solvent dielectric response[90] and also found an artificial minimum at large separation by using electrostatic cutoff methods as Dang and Pettitt,[89] in which the smooth truncation of electrostatics even amplifies this artifact. Alternately, the PMF profiles from simulations using Ewald sums produced the correct high dielectric shielding of the ions by water and removed the artificial minimum at large separation.

5. Conclusion

The current paper presents an extension of Ewald summation method to combined QM/MM calculations with semiempirical quantum models. The method is tailored to systems where the number of quantum atoms is small compared to the number of molecular mechanical atoms such that the Ewald contribution to the Fock matrix elements can be evaluated efficiently during the self-consistent field procedure required at each step of a molecular dynamics calculation. The method is based on a partition of the total Ewald potential into a short-ranged real-space interaction and a long-range periodic correction. The periodic correction term requires only a Mulliken charge representation of the charge density and hence can be used with any efficient linear-scaling Ewald method for point charge (or multipolar) systems, such as the particle mesh Ewald method. If the number of quantum atoms is sufficiently small, a considerable reduction in computational cost can be achieved through direct computation of the Ewald pair potential correction for only the quantum atoms such that the periodic correction to the electrostatic energy can be efficiently affected at each SCF iteration by a simple matrix multiplication with the Mulliken charge vector. Although the method is applied with semiempirical quantum methods in the present work, the methodology can be extended to other quantum models such as density-functional methods without significant code modifications. The implementation and performance of the method is tested in simulations of ion–ion association and on dissociative phosphoryl transfer reactions. The PMF profiles from these simulations are compared with those of full-electrostatic SBMD simulations and PBMD and SBMD simulations with electrostatic cutoff. Significant artifacts arise in the reaction free energies and activation barriers when a cutoff is used. These artifacts vanish when the QM/MM-Ewald method is employed. Despite the known problems associated with the use of electrostatic cutoffs, the majority of present day applications of QM/MM methods routinely employ cutoffs in simulations of biological reaction dynamics. Consequently, the present method offers an important extension of linear-scaling Ewald techniques to combined QM/MM calculations of large biological systems.

References

(1) Reichardt, C. *Solvents and solvent effects in organic chemistry*; VCH: Weinheim, 1990.

(2) Goedecker, S.; Scuseria, G. E. *IEEE Comput. Sci. Eng.* **2003**, *5*, 14−21.

(3) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.

(4) Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523−2544.

(5) Gao, J. *Rev. Comput. Chem.* **1995**, *7*, 119−185.

(6) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467−505.

(7) Warshel, A.; Åqvist, J. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 267−298.

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

(9) Warshel, A. *Acc. Chem. Res.* **2002**, *35*, 385−395.

(10) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186−195.

(11) York, D. M.; Darden, T.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *99*, 8345−8348.

(12) York, D. M.; Wlodawer, A.; Pedersen, L. G.; Darden, T. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8715−8718.

(13) York, D. M.; Yang, W.; Lee, H.; Darden, T.; Pedersen, L. G. *J. Am. Chem. Soc.* **1995**, *117*, 5001−5002.

(14) Sagui, C.; Darden, T. A. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155−179.

(15) Cheatham, T. E., III; Young, M. A. *Biopolymers* **2001**, *56*, 232−256.

(16) Makarov, V.; Pettitt, B. M.; Feig, M. *Acc. Chem. Res.* **2002**, *35*, 376−384.

(17) Grengard, L. *Science* **1994**, *265*, 909−904.

(18) Hockney, R. W.; Eastwood, J. W. *Computer Simulation Using Particles;* Institute of Physics Publishing: 1998.

(19) Gao, J.; Xia, X. *Science* **1992**, *258*, 631.

(20) York, D. M.; Lee, T.-S.; Yang, W. *J. Am. Chem. Soc.* **1996**, *118*, 10940−10941.

(21) Garcia-Viloca, M.; Truhlar, D. G.; Gao, J. *J. Mol. Biol.* **2003**, *327*, 549−560.

(22) de Leeuw, S. W.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. London, Ser. A* **1980**, *373*, 27−56.

(23) Ewald, P. P. *Ann. Phys.* **1921**, *64*, 253−268.

(24) Allen, M.; Tildesley, D. *Computer Simulation of Liquids*; Oxford University Press: 1987.

(25) Arfken, G. B.; Weber, H. J. *Mathematical methods for physicists*, 5th ed.; Academic Press: San Diego, 2001.

(26) Karasawa, N.; Goddard, W. A., III. *J. Phys. Chem.* **1989**, *93*, 7320−7327.

(27) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Hsing, L.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(28) Adams, D. J.; Dubey, G. S. *J. Comput. Phys.* **1987**, *72*, 156−176.

(29) Heyes, D. M. *J. Chem. Phys.* **1981**, *74*, 1924−1929.

(30) Perram, J. W.; Petersen, H. G.; Leeuw, S. W. D. *Mol. Phys.* **1988**, *65*, 875−893.

(31) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(32) Sagui, C.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2004**, *120*, 73−87.

(33) Bogusz, S.; Cheatham, T. E., III; Brooks, B. R. *J. Chem. Phys.* **1998**, *108*, 7070−7084.

(34) Boresch, S.; Steinhauser, O. *Ber. Bunsen-Ges. Phys. Chem.* **1997**, *101*, 1019−1029.

(35) Smith, P. E.; Pettitt, B. M. *J. Am. Chem. Soc.* **1991**, *113*, 6029−6037.

(36) Dewar, M. J.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899−4907.

(37) Dewar, M. J. S.; Thiel, W. *Theor. Chim. Acta* **1977**, *46*, 89−104.

(38) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833−1840.

(39) Löwdin, P. O. *J. Chem. Phys.* **1950**, *18*, 365−375.

(40) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735−746.

(41) Bader, R. F. W. *Acc. Chem. Res.* **1985**, *18*, 9−15.

(42) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 87−95.

(43) Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 1820−1831.

(44) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132−146.

(45) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **2003**, *24*, 1291−1304.

(46) Francl, M. M.; Carey, C.; Chirlian, L. E.; Gange, D. M. *J. Comput. Chem.* **1996**, *17*, 367−383.

(47) Sigfridsson, E.; Ryde, U. *J. Comput. Chem.* **1998**, *19*, 377−395.

(48) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700−733.

(49) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69−89.

(50) Dewar, M. J. S.; Zoebisch, E.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.

(51) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(52) Thiel, W. *MNDO97*, version 5.0; 1998.

(53) Stewart, J. J. P. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1−105.

(54) Thiel, W.; Voityuk, A. A. *J. Phys. Chem.* **1996**, *100*, 616−626.

(55) Gregersen, B. A.; Lopez, X.; York, D. M. *J. Am. Chem. Soc.* **2003**, *125*, 7178−7179.

(56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(57) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

(58) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384−2393.

(59) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697.

(60) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.

Efficient Linear-Scaling Ewald Method

J. Chem. Theory Comput., Vol. 1, No. 1, 2005 **13**

(61) Kumar, S.; Bouzida, D.; Swendsen, R.; Kollman, P.; Rosenberg, J. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(62) Brooks, C. L., III; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312−6325.

(63) Brooks, C. L., III; Brunger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843−865.

(64) Hutter, M. C.; Reimers, J. R.; Hush, N. S. *J. Phys. Chem. B* **1998**, *102*, 8080−8090.

(65) Schürer, G.; Gedeck, P.; Gottschalk, M.; Clark, T. *Int. J. Quantum Chem.* **1999**, *75*, 17−31.

(66) Bernal-Uruchurtu, M.; Ruiz-López, M. *Chem. Phys. Lett.* **2000**, *330*, 118−124.

(67) Bredow, T.; Geudtner, G.; Jug, K. *J. Comput. Chem.* **2001**, *22*, 861−887.

(68) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 498−510.

(69) Sherer, E. C.; York, D. M.; Cramer, C. J. *J. Comput. Chem.* **2003**, *24*, 57−67.

(70) Lopez, X.; York, D. M. *Theor. Chem. Acc.* **2003**, *109*, 149−159.

(71) Rozanska, X.; Chipot, C. *J. Chem. Phys.* **2000**, *112*, 9691−9694.

(72) Perreault, D. M.; Anslyn, E. V. *Angew. Chem., Int. Ed.* **1997**, *36*, 432−450.

(73) Florián, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 719−734.

(74) Florián, J.; Warshel, A. *J. Am. Chem. Soc.* **1997**, *119*, 5473−5474.

(75) Florián, J.; Åqvist, J.; Warshel, A. *J. Am. Chem. Soc.* **1998**, *120*, 11524−11525.

(76) Ahn, N. *Chem. Rev.* **2001**, *101*, 2207−2208.

(77) Guthrie, R. D.; Jencks, W. P. *Acc. Chem. Res.* **1989**, *22*, 343−349.

(78) Guthrie, J. P. *J. Am. Chem. Soc.* **1977**, *99*, 3991−4001.

(79) Ramond, T. M.; Davico, G. E.; Schwartz, R. L.; Lineberger, W. C. *J. Chem. Phys.* **2002**, *112*, 1158−1169.

(80) Range, K.; McGrath, M. J.; Lopez, X.; York, D. M. *J. Am. Chem. Soc.* **2004**, *126*, 1654−1665.

(81) López, C. S.; Faza, O. N.; Gregersen, B. A.; Lopez, X.; de Lera, A. R.; York, D. M. *Chem. Phys. Chem.* **2004**, *5*, 1045−1049.

(82) Gao, J. Computation of intermolecular interactions with a combined quantum mechanical and classical approach. In *Modeling the hydrogen bond*; Smith, D. A., Ed.; ACS Symposium Series 569; Oxford University Press: New York, 1994.

(83) Freindorf, M.; Gao, J. *J. Comput. Chem.* **1996**, *17*, 386-395.

(84) Riccardi, D.; Li, G.; Cui, Q. *J. Phys. Chem. B* **2004**, *108*, 6467−6478.

(85) *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P., Mallard, W., Eds.; National Institute of Standards and Technology: Gaithersburg MD, 20899, 2003 (http://webbook.nist.gov).

(86) Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77*, 334−338.

(87) Jencks, W. P. *Chem. Rev.* **1985**, *85*, 511−527.

(88) Bader, J. S.; Chandler, D. *J. Phys. Chem.* **1992**, *96*, 6423−6427.

(89) Dang, L. X.; Pettitt, B. M. *J. Phys. Chem.* **1990**, *94*, 4303−4308.

(90) Del Buono, G. S.; Figueirido, F. E.; Levy, R. M. *Chem. Phys. Lett.* **1996**, *263*, 521−529.