Annu. Rev. Biophys. Biomol. Struct. 1999. 28:155–79

MOLECULAR DYNAMICS SIMULATIONS OF BIOMOLECULES: Long-Range Electrostatic Effects

Celeste Sagui and Thomas A. Darden
National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709; e-mail: darden@niehs.nih.gov

KEY WORDS: free energies, particle mesh Ewald, fast multipole, periodic boundary conditions, Ewald summation

ABSTRACT

Current computer simulations of biomolecules typically make use of classical molecular dynamics methods, as a very large number (tens to hundreds of thousands) of atoms are involved over timescales of many nanoseconds. The methodology for treating short-range bonded and van der Waals interactions has matured. However, long-range electrostatic interactions still represent a bottleneck in simulations. In this article, we introduce the basic issues for an accurate representation of the relevant electrostatic interactions. In spite of the huge computational time demanded by most biomolecular systems, it is no longer necessary to resort to uncontrolled approximations such as the use of cutoffs. In particular, we discuss the Ewald summation methods, the fast particle mesh methods, and the fast multipole methods. We also review recent efforts to understand the role of boundary conditions in systems with long-range interactions, and conclude with a short perspective on future trends.

CONTENTS

1056-8700/99/0610-0155$08.00

INTRODUCTION

The explosive growth of computer power over the past two decades has led to the development of large-scale simulation techniques whose aim is to directly reproduce or simulate processes on a molecular level. Molecular dynamics (MD) simulations, of both a classical and quantum nature, have proven to be invaluable in elucidating the structural, mechanical, electrical, and chemical properties of diverse sets of materials. For example, MD simulations have been used to study the liquid state, the bulk solid, diffusion, wetting phenomena, phase transformations, polymers, and protein dynamics, to name several examples. Indeed, it may be argued that MD simulations [and variants (1)] represent the future of theoretical endeavors in the fields of physics, chemistry, and molecular biology.

The simulation of biologically active molecules poses its own unique set of problems to the computational scientist. When contemplating an MD simulation of biomolecules, one would ideally like to carry out a quantum mechanical calculation, based for example on the density functional theory approach (DFT). Schemes such as the Fast-Fourier transform based Car-Parrinello technique (2) or other real-space multigrid methods (3) have reached maturity. These methods are known to be reliable, with truly predictive powers. Indeed, they have provided us with some of the most accurate theoretical descriptions of materials to date. However, they require substantial computational investments. With state-of-the-art parallel supercomputers, it is now possible to simulate hundreds of atoms over picosecond timescales. Clearly, such simulations are as yet impractical for biomolecular systems, which typically contain many thousands of atoms with relevant timescales of nanoseconds to seconds. Hence, classical MD methodology is the technique of choice for simulating biomolecules. With current technology, classical simulations of systems consisting of tens of thousands of atoms and with running times of tens of nanoseconds are rapidly becoming common.

In classical MD simulations, the trajectories of a group of interacting atoms are calculated through a discretization of Newton's laws. Forces are generated by atom-atom interactions, which are usually given in terms of an empirical potential. Parameters for the potential are usually obtained by fitting to either ab initio calculations or experimental data. A typical empirical force field

formulation of the potential energy for N atoms at positions $\mathbf{r}_1, \ldots, \mathbf{r}_N$, can be written as:

$$E = \sum_{\text{bonds}} \frac{a_i}{2}(l_i - l_{i0})^2 + \sum_{\text{angles}} \frac{b_i}{2}(\theta_i - \theta_{i0})^2 + \sum_{\text{torsions}} \frac{V_n}{2}(1 + \cos(n\omega - \gamma))$$

$$+ \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{q_i q_j}{r_{ij}}, \qquad 1.$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ (4). In this expression, the first three terms deal with the specific internal degrees of freedom within molecules. The first term is a harmonic potential between bonded atoms that gives the contribution to the energy when the bond length l_i deviates from the equilibrium value l_{i0}. Similarly, the second term is a harmonic potential in the valence angles of the molecules. The third term is a torsional potential describing the periodic variation in energy due to bond rotations. Low-order Fourier series in the dihedral angle ω are used, where V_n represents the barrier height and γ the phase shift of the n-fold term. The last two terms represent nonbonded interactions. The fourth contribution is a Lennard-Jones potential representing the van der Waals interactions, and the last term is the Coulomb electrostatic potential. Note that the first four terms deal with mainly short-ranged interactions. There are successful methods for treating these short-range terms, although more sophisticated and computationally challenging formulations will be required in the future (5–7). These short-range interactions will not be discussed here. Rather, we will concentrate on the treatment of the long-range electrostatic potentials, which is where the current challenges lie.

The need for a correct treatment of long-range electrostatic forces in simulations of biomolecular systems has been clearly established in the last decade. Historically, long-range forces were mainly ignored in macromolecular simulations, being truncated with the use of artificial, nonbonded cutoffs. The reasons for this were basically twofold. First, there was insufficient computer power available to perform the $O(N^2)$ calculation necessary to evaluate all nonbonded pairs in macromolecular systems. Second, while all would agree that full pair evaluation is correct for isolated systems, many workers have felt that full evaluation of Coulombic interactions in periodic boundary conditions would lead to substantial artifacts when simulating solution phase systems. Improvements in computer speed together with improved algorithms have now eliminated the first objection to including long-range interactions. These can now be included at a cost not much more than that of a typical cutoff calculation. The use of these methods has led to radically improved results in many cases, particularly with

respect to nucleic acid simulations. These developments are reviewed more fully by Auffinger & Westhof (8) and Cheatham et al (9).

The use of cutoffs has been shown to lead to severe artifacts in simulations of peptides and nucleic acids (10, 11). Although the most sophisticated cutoff methods (12) can yield stable DNA simulations, at least on the nanosecond time scale (13), even these have been shown to fail for interfacial and membrane simulations (14). On the other hand, the second objection above to the use of full electrostatics in periodic boundary conditions with highly polar or charged systems continues to be raised (15). In liquid simulations, long-ranged electrostatics have long been treated fully, using Ewald summation, and periodicity artifacts were shown to be minimal in this context. Recently, several groups have examined the artifacts due to Ewald summation in charged or polar systems of biological relevance. The results so far point to minimal periodicity artifacts, at least for systems in high dielectric solvents. Below we discuss these results more completely.

This article is organized as follows: In the next section, we introduce the basic issues involved in accurately representing the electrostatic potential and forces due to a charge distribution, pointing out the physical problems associated with finite system size, and the formidable computational problems due to the size of relevant biomolecular systems. Next, we discuss Ewald summation and the fast particle mesh–based approximations to it. In the next section, we discuss the fast multipole method (FMM) and related tree algorithms. Following this, we review current efforts to understand the effect of long-range boundary conditions on simulation results. Finally, we offer a concluding perspective.

ELECTROSTATICS: A MATHEMATICAL AND COMPUTATIONAL PROBLEM

In this section we consider specifically the electrostatic interactions in a system of particles. These particles can be atoms, small molecules (e.g. water), protein residues, etc. We assume that all nonelectrostatic forces between the particles, such as chemical bonding and attractive van der Waals forces, are known. In principle, these may be computed by fitting a model potential to experimental or ab initio data. To complete the entire physical description of the problem, we therefore need to concern ourselves primarily with the long-range electrostatic interactions between the particles. We do not attempt to describe the long-range induction interactions here, although these are of increasing interest in biomolecular simulations. However, we note that calculation of induction effects depends on a correct description of the electrostatic interactions.

Intrinsic to the electrostatic problem is the question of how charge may be distributed in space and how such a distribution is described mathematically.

A charge distribution can be either continuous or discrete. At the most fundamental level, quantum mechanics teaches us that while the positive nuclear charge may be considered discrete on the atomic scale, the negative electronic charge is distributed continuously in the electronic clouds or orbitals, as dictated by the solutions of the electronic Schroëdinger equation. Since the size of biomolecules generally precludes quantum mechanical molecular simulations, this basic continuous distribution is forsaken in favor of a discrete set of point charges, dipoles, and perhaps higher-order multipoles, all characterized by vanishingly small dimensions. The total electrostatic potential ϕ due to the charge distribution is then expressed as a sum of successive multipolar potentials: a monopolar potential ϕ_0, a dipolar potential ϕ_1, a quadrupolar potential ϕ_2, etc. In principle, such a "distributed multipole" description can exactly describe the potential ϕ due to the true charge density, at points distant from the expansion centers where "penetration" effects are negligible (7). In practice this expansion is truncated, usually at low order, and often at the monopole level in current force fields. The distributed monopoles are usually referred to as partial charges, since they need not have integral values. The long-range effects are most pronounced for the monopolar potential ϕ_0, and thus, for the rest of this review, only the monopolar potential is considered; the extension to higher-order multipolar interactions proceeds along similar lines to the treatment of the monopolar term; see, for instance, the treatment of point multipoles in the context of Ewald summations by W Smith (16).

The energy of a system of N partial charges q_i, that produce the monopolar electrostatic potential ϕ_0 is given by:

$$E(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{1}{2}\sum_i^N q_i \phi_0(\mathbf{r}_i) = \frac{1}{2}\sum_i^N \sum_{j \neq i}^N \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|}, \qquad 2.$$

where $\phi_0(\mathbf{r}_i)$ is the monopolar potential acting on charge i at position $\mathbf{r}_i$, produced by all the other charges q_j, with $j \neq i$. Physically, Equation 2 represents a well-posed electrostatic problem: By knowing all the partial charges q_i and their positions $\mathbf{r}_i$, one can compute the electrostatic interactions, no matter how complicated the configurations of charges may be. In practice, such computation is far from trivial, due to two main computational considerations:

1. Finite system size: The treatment of infinitely long-range interactions in a sample of finite size is a long-standing technical issue. Finite-size constraints being inescapable, the infinite range of the interactions, as well as other quantities that characterize the properties of the system in the thermodynamic limit, are dealt with through boundary conditions.

2. Molecular size: Biomolecules (e.g. DNA strands, proteins, membranes, enzymes, etc.) range in size from a few tens to millions of atoms. System sizes currently simulated with MD techniques range from 10^3 to 10^4 atoms, excepting the rare 10^5 atom simulation. Equation 2 represents a sum over $N(N-1)/2$ pairs; i.e. it is an $O(N^2)$ calculation, which becomes too costly as N grows larger than 10^3.

Boundary conditions may be conveniently divided into periodic boundary conditions (PBC) and nonperiodic boundary conditions (NPBC). The latter include vacuum simulations, in which solvent is ignored completely, and the so-called continuum methods, which treat the solvent implicitly. They also include intermediate approaches, models with minimal explicit solvation, in which the solute is surrounded by a thin layer of explicit waters. On the other hand, PBC mainly include simulations involving explicit solvent molecules (though this is not always the case; for instance, the reaction field methods use PBC while treating neighboring solvent explicitly and distant solvent through continuum methods). Although both NPBC and PBC have been criticized, because they can introduce serious artifacts in the simulations, there is growing consensus on the soundness of PBC in the treatment of long-range forces. We review recent work on this problem in the section on artifacts due to Ewald summations.

With respect to the second problem, two major approaches have proved to give satisfactory and consistent reproduction of the electrostatic interactions, as well as feasible computer times:

1. Ewald summation and other lattice summation techniques: In periodic boundary conditions, these methods consider the potential due to the partial charges of a system, together with all of their periodic images. By using the decomposition $1/r = \mathrm{erfc}(\beta r)/r + \mathrm{erf}(\beta r)/r$, one splits the Coulomb interaction into a short-range term, handled exactly in the direct sum, plus a long-range, smoothly varying term, handled approximately in the reciprocal sum by means of Fourier methods. Note that important, close-range interactions such as those in hydrogen-bond pairs are partially calculated in the approximate reciprocal sum, which is a possible limitation of these methods, particularly with regard to multiple time-step implementations. However, the splitting of the Coulombic interactions involves a smooth function of r and thus, these approaches yield very well behaved MD simulations. These methods include the original Ewald summation (17), the particle-particle particle-mesh method (18, 19), and the particle-mesh Ewald algorithm (20, 21).

2. Fast multipole methods: These methods treat Coulombic interactions exactly for particles within the same or neighboring subcells of the simulation cell, and evaluate the potential for more distant particles through multipolar

expansions. Note that close-range interactions are handled exactly, unlike those in the lattice sums. However, the splitting of the Coulombic interactions is not a smooth function of r, and thus, these methods may suffer in comparison to lattice sum methods in their simulation behavior (22). While the original FMM (23, 24) uses NPBC; more recently it has also been implemented for PBC (19, 25).

We review both approaches in the next two sections.

EWALD SUMMATION AND OTHER LATTICE SUMMATION TECHNIQUES

Two Approaches to Ewald Summation

In PBC, the system to be simulated comprises a unit simulation cell U, whose edges are given by the vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$, not necessarily orthogonal. The volume of U is given by $V = \mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3$. The Coulombic potential under PBC can be treated in two similar, but not identical ways. (We ignore cutoff and reaction field treatments, which do not explicitly treat electrostatics with full periodicity.) The first way considers that the potential due to the charges is given by the solution of Poisson's equation in PBC. For simplicity, assume that there is a single charge q in the unit cell, situated at the origin, and the potential $\phi(\mathbf{r})$ is desired at another point $\mathbf{r} \neq \mathbf{0}$. To arrive at a tractable expression (26, 27), a smoothly varying screening charge distribution (typically a Gaussian), centered at the origin, is added to the point charge q such that the total charge of this cloud (given by the integral of the distribution over all space) exactly cancels q. Thus, the electrostatic potential at $\mathbf{r}$ is produced entirely by the fraction of q that is not screened. This fraction decays very rapidly with distance. Because a screening charge cloud has been added to the point charge, one must add an exactly compensating charge distribution. This second distribution varies smoothly in space and acts as a source term for the solution of Poisson's equation in PBC, which can be expressed as a rapidly convergent Fourier series. However, since the solution for Poisson's equation in PBC is undefined for non-neutral unit cells (otherwise, the zeroeth-order term for the potential would be infinite), a uniform density having total charge $-q$, referred to as a uniform neutralizing plasma, is added to the source term. This fixes the problem with the zeroeth-order term, without disturbing the higher-order terms. This approach to solving the Poisson problem gives $\phi(\mathbf{r})$ for $\mathbf{r} \neq \mathbf{0}$. The potential at the origin is obtained by removing the direct Coulombic contribution before taking the limit $|\mathbf{r}| \to 0$. This defines the Wigner or self potential ζ:

$$\zeta = \lim_{|\mathbf{r}|\to 0} \left(\phi(\mathbf{r}) - \frac{1}{r} \right). \qquad 3.1$$

In cubic unit cells, where $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are mutually orthogonal with length L, ζ is given by $\zeta \approx -2.837297/L$. The potential at $\mathbf{r}$ due to the charge q_i at $\mathbf{r}_i$ is given by $q_i\phi(\mathbf{r}-\mathbf{r}_i)$, and the energy of the unit cell consisting of point charges $q_1, \ldots, q_N$ at positions $\mathbf{r}_1, \ldots, \mathbf{r}_N$ is given by

$$E = \frac{1}{2}\sum_{i\neq j} q_i q_j \phi(\mathbf{r}_i - \mathbf{r}_j) + \frac{\zeta}{2}\sum_{i=1}^{N} q_i^2. \qquad 3.2$$

Forces are obtained by differentiating the energy with respect to particle positions, and the virial part of the pressure tensor is obtained by differentiating with respect to the cell vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ (28). As noted above, the Ewald sum may readily be extended to multipolar interactions. In molecular systems there is an additional consideration. Bonded atoms, such as atoms belonging to the same water molecule, do not usually interact electrostatically and thus should not feel the full potential ϕ. Typically the direct Coulomb interaction is removed from the periodic potential ϕ, analogous to the process leading to the self potential ζ.

The second approach (29–31) to treat Coulombic interactions in PBC is to model the system as a large but finite array of copies of the unit cell U, which is then immersed in a dielectric medium. Each particle i at position $\mathbf{r}_i$ within the cell has a number of image particles at positions $\mathbf{r}_i + \mathbf{n}$, with $\mathbf{n} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$ and n_1, n_2, n_3 integers. The interaction of particle i with j in periodic boundary conditions is evaluated by summing the interaction of i with j and with all images of j. Particle i also interacts with its own periodic images. Finally, particle i interacts with the reaction field induced in the surrounding dielectric medium. In the limit as the array size tends to infinity, this implementation of the PBC leads to infinite lattice sums, and the monopole potential $\phi_0(\mathbf{r}_i)$ acting on charge i due to the array of copies of U is given by

$$\phi_0(\mathbf{r}_i) = \sum_{\mathbf{n}}{}' \sum_{j=1}^{N} \frac{q_j}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{n}|}, \qquad 3.3$$

where the prime indicates that terms with $i = j$ and $\mathbf{n} = 0$ are omitted.

Unfortunately, the Coulomb interactions do not decay fast enough to provide for an absolutely converging series. To appreciate the problems involved, consider the simple lattice sum

$$\sum_{\mathbf{n}\neq 0} \frac{1}{|\mathbf{n}|^p}, \qquad 3.4$$

where the lattice vectors $\mathbf{n}$ are defined as above. If $p > 3$, this sum converges, whereas if $p \leq 3$, it diverges. This result is the origin of the distinction between

short-range and long-range interactions. Thus, if there were only a single positive charge in the unit cell, replicated by periodic boundary conditions, the unit cell electrostatic energy would be infinite even though, by symmetry, the particle would experience no force. Similarly, if the unit cell were not neutral, the sum in Equation 3.3 would diverge. Oddly enough, the forces actually converge (conditionally) in this case, although they are long ranged as well.

If the unit cell is neutral, the sum in Equation 3.3 converges, albeit only conditionally. This latter statement means that the result obtained depends on how the limit to infinity is taken. A simple example of a conditionally convergent series is the one-dimensional series $1 - 1/2 + 1/3 - 1/4 + 1/5 + \cdots$. The series converges, meaning the sequence of its partial sums converges to some number. However, if the series is rearranged as follows: $1 + 1/3 - 1/2 + 1/5 + 1/7 - 1/4 + \cdots$, that is, two positive terms followed by a negative term, then it can be shown (32) that it now converges to a different limit. This peculiar result cannot happen if a series is absolutely convergent, that is, if the series with all terms replaced by their absolute values were to converge. The same kind of problem can be expected for the sum in Equation 3.3, since it is not in general absolutely convergent, by the above criterion Equation 3.4, i.e. Coulombic interactions are long ranged. If, for example, you perform the summation by slabs, making n_3 the outer sum, and first taking the infinite sum over n_1 and n_2 for each fixed n_3, you will not, in general, get the same answer as if you compute the sum by gathering the lattice points **n** into spherical shells and summing this way. The forces will not agree either.

Ewald first tackled the problem of long-range interactions between particles and their infinite periodic images in 1921 (17). He transformed the slowly, conditionally convergent sum for the Coulomb potential into two sums that converge rapidly and absolutely: the direct and the reciprocal sums. The Ewald sum agrees with the result in Equation 3.2. Note that a conditionally convergent sum cannot be transformed into an absolutely convergent sum without losing something. DeLeeuw et al (29) showed that the lattice sum in Equation 3.3 can be re-expressed as the Ewald sum in Equation 3.2 plus a term that depends on the dipole moment of the unit cell U, the dielectric constant of the surrounding medium, and the order of summation; that is, on the asymptotic shape of the array of copies of U. Thus, the second way of treating electrostatics in the context of PBC does not yield identical results to the first. If the surrounding medium has infinite dielectric (tin-foil boundary conditions), the extra term in the treatment of DeLeeuw et al vanishes, leaving the Ewald sum for energies and forces. However, Smith (30) showed that the pressure tensor is different from that for Ewald summation, even in tin-foil boundary conditions. The above second term of DeLeeuw et al is only rarely used, however; Boresch & Steinhauser have recommended its use in calculations of the dielectric constant

of the sample (33). Note that this second term is computationally inexpensive, compared with the rest of the Ewald sum.

Note that with the second approach, the self-energy ζ has a natural interpretation. It is the potential acting on a charge due to its own periodic images, together with the neutralizing plasma. A similar interpretation can be given to the corrections for bonded pairs in molecules. Thus, the second approach has a more physical interpretation in terms of crystals, at least for static configurations. However, the array of periodic images of atoms must be perfectly correlated dynamically, which does not correspond to a real crystal. Recently, this second approach to Coulombic interactions in PBC has been used to formulate the quantum mechanical Hamiltonian in PBC (34). It is possible that ab initio treatment of crystals and other condensed-matter systems may provide a test to decide which of the two ways of treating electrostatics in PBC is more physically sound.

The Ewald direct sum is given by the sum of screened interactions, using Gaussian screening functions centered at each point charge. The reciprocal sum is given by the Fourier series representation of the solution to Poisson's equation in PBC, using the sum of the compensating Gaussians, again centered at each point charge, as a source distribution. The terms in this Fourier series can be expressed analytically. By varying the width of these Gaussian distributions, the relative contributions and computational cost of the direct and reciprocal sum can be varied without affecting the total energy, forces, etc. If the Gaussians are chosen to vanish (within a prescribed tolerance) at a cutoff of half the cell size, the direct sum is over all minimum image pairs, and thus is $O(N^2)$, while the number of terms needed in the reciprocal sum is $O(1)$, and thus the reciprocal sum is $O(N)$. If, on the other hand, the width is taken so that the Gaussian vanishes at a standard cutoff distance independent of N, conventionally taken to coincide with the cutoff of the Lennard-Jones interactions, the direct sum is $O(N)$, but the number of terms needed in the reciprocal sum is $O(N)$, so that the reciprocal sum becomes $O(N^2)$ (21). By varying the cutoff with the square root of the cell size, it can be shown that both the direct and reciprocal sum are $O(N^{3/2})$, which is optimal (35).

Thus, the Ewald sum, while immensely faster than explicit evaluation of the lattice sum in Equation 3.3, is computationally expensive for large systems. This cost can be alleviated by replacing the Gaussian screening functions by an optimal screening function (36) or by multiple time-step approaches wherein local Coulomb interactions are calculated every time step, with full Ewald performed only every several steps (37). The latter approach, using the r-RESPA methodology (38), provides a four- or more-fold speedup for large systems while conserving energy and reproducing the density of states. However, even with these optimizations, Ewald summation remains costly compared to conventional cutoff schemes.

Particle Mesh–Based Approaches to Lattice Summation

The particle mesh–based approaches all attempt to accelerate the solution of Poisson's equation in PBC, using the profound advantages of the fast Fourier transform (FFT) for calculating discrete Fourier transforms. As stated above, when a fixed cutoff is applied to the direct Ewald sum, the number of terms needed in the reciprocal sum is $O(N)$. Since each such term is nominally of order N to calculate, the reciprocal sum is $O(N^2)$. Similarly, the discrete Fourier transform of N coefficients is an order N^2 calculation. However, the FFT performs this task in $O(N \log N)$ operations. All the particle mesh—based approaches reduce the calculation of the reciprocal sum to a sum over coefficients of the discrete Fourier transform of a gridded charge distribution, which is then accelerated to an $O(N \log N)$ calculation using the three-dimensional FFT. The methods differ in how they transform the continuous charge density due to the sum of compensating Gaussians onto a regular three-dimensional grid and in how they compensate for the loss of accuracy introduced in this process. These methods have been carefully studied and compared in the recent literature (19, 39–41). Here we attempt to briefly review the similarities and differences of the different variants of the particle-mesh method, pointing out the strengths and weaknesses of each.

Hockney & Eastwood (18) first developed the particle-mesh method, wherein the Coulombic potential at particles was obtained by solving Poisson's equation in periodic boundary conditions. The charges were interpolated onto a regular grid, and the discretized Poisson's equation was solved by expanding in a discrete Fourier transform. For a regular grid, this transform is evaluated using the FFT, resulting in a very efficient algorithm. However, the interactions between nearby particles were poorly represented. Subsequently they developed the particle-particle particle-mesh approach (PPPM or P3M), wherein the interaction is split into short- and long-range contributions using switching functions and the long-range potential is obtained by gridding the charges and employing the FFT as above. Note that the switching function, analogous to the Gaussian screening and compensating distributions in Ewald summation, leads to a smoother charge density to interpolate onto grids. Higher accuracy is achieved by using an optimal Green's function, obtained through a least-squares approach. A readable account of this least-squares optimization of the Green's function is given by Ferrell & Bertschinger (42).

As noted above, the original particle-mesh method, although appropriate for collisionless systems such as plasmas, was not accurate enough for MD of molecular systems (43, 44). Furthermore, in their published work on P3M, Hockney & Eastwood discuss only low-order implementations of their method, which again are insufficiently accurate for molecular simulations. In

addition, the connection with the well-understood Ewald summation was not clear, probably because the authors did not apply Gaussian screening functions. Indeed, Ewald summation is hardly mentioned in their book. Probably for these reasons, the particle-mesh–related approaches were overlooked by the molecular simulation community for many years in favor of the multipole expansion approaches such as FMM (44). Recently, variants of this approach have been taken up by various groups (19–21, 45, 46), and it currently appears to be the most popular approach to efficiently approximate the Ewald sum.

Assuming the usual Gaussian screening function, leading to the Ewald direct and reciprocal sum, the direct sum is handled precisely as in the Ewald sum, i.e. usually along with the other nonbond terms, using a Verlet list of neighbors. The steps in the particle-mesh algorithm for the reciprocal sum are as follows:

1. Charge assignment: Charge is smoothly interpolated to neighboring grid points, using a weighting function. The cost of this step is $O(Np^3)$, where p is the order of the weighting, e.g. cubic interpolation is order 4.

2. Grid transformation: The Fourier space representation of the discrete charge distribution is obtained via the FFT. The cost of this step is $O(K \log K)$, where K is the number of grid points. For condensed-matter systems with density near that of water, K is linearly proportional to N.

3. Multiplication by the optimized influence function: The components of the transformed charge grid are multiplied by a type of Green's function that, in the absence of optimization, would be the normal expression in the reciprocal Ewald sum. At this point, forces may be calculated in Fourier space, as in the Ewald sum. This approach (referred to as force interpolation), which leads to three grids to transform back by FFTs, is favored by Pollock & Glosli (19). The original approach, which was implemented by Shimada et al (45) and Luty et al (46), fills only one grid. In either case, the cost of this step is $O(K)$.

4. Transformation back to real space: Either one or three grids are transformed back. The cost of this step is again $O(K \log K)$.

5. Force assignment: In the force interpolated case (19), the three components of the force from the three transformed grids are interpolated onto the atoms, using the same weighting functions as in step 1. Alternately, the force at grid points is approximated by finite differencing the potential on the transformed grid (45, 46). The three components of the approximate force are then interpolated onto atoms as in the force interpolation case. In either case, the cost of this step is $O(Np^3)$.

Thus, in either variant of P3M, the overall cost is $O(N \log N)$, since K is proportional to N. The finite difference approach to forces requires fewer FFTs (two versus four); however, Deserno & Holm (41) showed that it is significantly less accurate than the force interpolation approach.

The PME algorithm (20, 21) approaches the problem from a very different philosophy, although the resulting algorithm is very similar to that of P3M. In the PME, the basic form of the Ewald sum is taken as a given. Rather than approximating the Gaussian source density by the weighting function, interpolating charge onto grid points, and then adjusting the Green's function, the complex exponentials appearing in the reciprocal sum are approximated by local polynomial interpolation. In the original implementation (20), local Lagrange interpolation was used. In the smooth PME (21), the Euler spline interpolation, based on B-splines, was used. The original implementation used force interpolation, resulting in four total FFTs. The smooth PME took advantage of the smoothness of the B-splines, together with recursion formulas for their analytic derivatives, to arrive at forces through analytic differentiation in real space, using only two FFTs. The actual steps in the algorithm are quite close to the P3M outlined above and can be described using the same five steps. The smooth PME differs in the last step in that the potential is first interpolated onto the atom, using the Euler spline weighting functions, and that the resulting potential is differentiated with respect to atomic position to give the electric field and hence the forces.

The weighting functions used in the P3M are also B-splines, although that does not seem to have been noted until recently. Thus the P3M is even closer to the PME than previously noted. Darden et al (40) compared the accuracy of the smooth PME and a force-interpolated variant of the smooth PME (i.e. forces obtained by interpolating the Ewald forces, but using Euler splines), with the force-interpolated version of P3M implemented by Pollock & Glosli (19), using a program kindly provided by E Pollock. They found that the Pollock implementation was more accurate than the smooth PME or its force-interpolated variant. This difference in accuracy, when testing actual molecular systems, was found to be marginal for cutoffs and overall accuracies typically employed in simulations (e.g. 9 Å cutoff, RMS force error of $\sim 10^{-4}$), and in these cases the smooth PME (forces via analytic differentiation) could achieve equal or higher accuracy by using one higher order of interpolation. Since the smooth PME uses half as many FFTs as the force-interpolated P3M, it is more efficient in situations where the cost of the FFT becomes an issue, such as in highly parallel implementations (the cost dependence on spline order is in steps 1 and 5 above, which parallelize well, whereas the FFT does not). Nonetheless, it was of interest to trace the difference in accuracy between the PME and P3M.

Note that the optimal P3M influence function is obtained by assuming that charges are randomly and independently distributed in the cell. This suggested that the smooth PME could be improved by using least-squares spline approximation of the complex exponentials, rather than spline interpolation. Since there is an exact Fourier series representation of the Euler spline, the coefficients of the least-squares approximation can be obtained analytically using the well-known orthogonality properties of Fourier series (40). The resulting optimized force-interpolated PME gave the same accuracy as the Pollock & Glosli implementation. However, the resulting Green's function is different (with PME computationally cheaper, although this cost is minor) than that in P3M. The difference between the two is that the P3M attempts a least-squares fit to the exact Ewald sum, whereas the optimal PME can be re-expressed as a least-squares fit to the truncated Ewald sum, expressed as the sum over the reciprocal vectors on the grid. The two give the same accuracy because the error in the Ewald sum due to this truncation is very small compared to the error introduced by approximating the terms in the truncated series. That is, the truncated Ewald sum is far closer to the exact Ewald sum than the particle-mesh approximations are to either.

Later, Deserno & Holm (41, 41a) have more thoroughly compared the PME methods with the above two implementations of the P3M, although they did not consider the optimal force-interpolated PME in their comparison. Otherwise, their conclusions are similar. The most accurate algorithm is the force-interpolated P3M. The finite-difference implementations of P3M also use half as many FFTs as the force-interpolated variant and thus enjoy the same cost advantages as the smooth PME, but Deserno & Holm showed they are less accurate than the smooth PME, which uses analytic differentiation rather than the finite-difference approximation to arrive at forces. Thus, the overall conclusions from these two studies would be that the force-interpolated P3M and the optimized force-interpolated PME are the most accurate for a given grid size and spline order. In cases where it is less expensive to use a higher-order spline than to perform the extra two FFT's, the smooth PME is preferable. The cost advantages of using the analytic differentiation approach to forces and fields will increase when force fields involving point dipoles and higher-order multipoles are considered.

Another variant of the particle-mesh approach is the fast Fourier Poisson method, proposed by York & Yang (47). This method directly samples the compensating Gaussians onto the grid and avoids loss of accuracy from interpolation by use of a clever identity. Due to the cost of sampling the Gaussians, this method is not competitive for the modest accuracies appropriate for current force fields, but it appears to be more efficient than the above methods for high-accuracy requirements that might be expected in future ab initio calculations.

FAST MULTIPOLE METHOD

Description of the Method

In this section we review the fast multipole method of Greengard and Rokhlin (23, 44, 48, 49). The method can be discussed in terms of a number of contributing developments.

MULTIPOLE EXPANSIONS The FMM relies on standard multipole expansions for the electrostatic potential as described, for instance, in Jackson (50). The multipoles are expressed in terms of spherical coordinates (r, θ, α) rather than Cartesian coordinates, since expansions in the spherical harmonics $Y_{lm}(\theta, \alpha)$ are far more efficient at higher order. If the distance to charge q_i is $r_i = |\mathbf{r}_i|$, the radial expansion is performed in powers r_i^l / r^{l+1} ("multipolar" expansion) when $r > \max\{r_i\}$ and in powers $r^l / {r_i}^{l+1}$ ("local" expansion) when $r < \min\{r_i\}$. It can be shown that the infinite series in l in the multipole expansion can be truncated with a cut-off p to an accuracy $\frac{C}{r - r_i^{\max}} (\frac{r_i^{\max}}{r})^{p+1}$, with $C = \sum_{i=1}^{N} |q_i|$.

OCTAL TREE CONSTRUCTIONS There are several previous methods, known as tree algorithms, that are based on the recursive grouping of distant particles into multipoles. These methods, as first introduced by Appel (51) and Barnes & Hut (52), are based on a hierarchical tree approach to calculating the energies and forces in a system of N particles. Although many of these algorithms were originally developed for astronomical simulations (53), where the monopole term plays an important role, they are easily modified for a system of charges, where higher-order expansions are necessary due to approximate local charge neutrality. The FMM uses a similar hierarchical tree construction.

The tree method starts by successively dividing the initial cell into self-similar subcells, known as children. Thus, if the initial cell (level zero) is a cube, the first division produces eight equal cubes (level 1), the second division produces 64 cubes (level 2), and the k division produces 8^k cubes. The cells at level l are the parents of cells at level $l+1$ and the children of cells at level $l-1$. The cells at the finest subdivision level are called the leaves. Generally, the refinement process stops at $\sim \log_8 N$ number of levels, which means that there are $O(N)$ leaves containing $O(1)$ particles. The FMM allows a larger number of particles in the leaf cells, to optimize efficiency.

Two cells at the same level are nearest neighbors if they share a boundary point, and are well separated when they are not nearest neighbors. Obviously at level 0 and 1, all cells are nearest neighbors. For each cell i at level l, there is an interaction list formed by other cells at level l that are the children of the nearest neighbors of i's parent and are well separated from i. For each cell i at

level l, a multipole expansion is performed about the cell center. The expansion represents the far field created by the particles inside the cell.

The key of the tree methods is to compute interactions between well-separated clusters through multipole expansions with rigorous bounds on the error. For a given precision ϵ, it is necessary to use $p = -\log_{\sqrt{3}}(\epsilon)$ terms in the expansion (49). The interactions between particles contained in each cell with those contained in well-separated cells is computed recursively by evaluating the multipole expansion corresponding to cell i at the position of every particle in the interaction list of cell i.

At this stage, a tree algorithm is of order $O(N \log N)$. Each particle contributes to p^2 expansion coefficients, so that all expansions are created in Np^2 operations. For each particle, the maximum size of the interaction list is 189. Thus, the total number of evaluations is $189Np^2$. This must be repeated at each tree level, resulting in $189Np^2 \log_8 N$ evaluations. At the finest level, there are nearly N leaves with $O(1)$ particles each, so that the interactions between nearest neighbors, which are calculated without approximation, require $27N$ operations. The total cost is approximately $189Np^2 \log_8 N + 27N$. As it stands, the method is still very costly; for high precision and $N \sim 10^5$, the order of expansion $p \sim 20$ is so big that the method is only two to three times faster than the direct Coulomb sum (49).

TRANSLATION OPERATORS FOR TRANSFORMATION OF MULTIPOLE EXPANSIONS

The FMM reduces the asymptotic cost of the previous tree algorithm through three translation operators: a translation of a multipole expansion, a conversion of a multipole expansion into a local expansion, and a translation of a local expansion. The implementation of the original FMM involves a two pass procedure.

In the first or upwards pass, multipole moments are computed for every leaf cell. For all the other levels, the multipole translation operator is used to shift the center of expansion of a multipole moment from a cell's center to that of its parent. The contributions of the eight children, weighted by the translation operator coefficients, are summed together to give the multipole moments of the parent.

In the second or downward pass, the expansion at level l for cell i (with l going downwards from level 2 to the leaves), is given by the sum of two contributions. With the use of the conversion operator, the multipole expansion corresponding to each cell in the interaction list of cell i is transformed into a local expansion around the center of cell i. The sum of the converted local expansions due to all the members of the interaction list gives the first term in the potential of cell i. The second term is obtained by shifting the local expansion of i's parent to i itself, with the use of the local operator.

Finally, for each particle in each leaf cell, one evaluates the local expansion potential at the particle position and computes the interactions with particles in

near neighbor cells directly. (Their contribution is not in the local expansion coefficients, because they do not belong to the interaction list.)

An evaluation of the number of operations in this implementation of the FMM gives an estimate that apparently would beat that of the other tree algorithms. Unfortunately, the order of expansion p needed for a given precision in the latter is smaller than that needed in the FMM. The major hurdle in the FMM scheme is the conversion of multipole to local translations in the downward pass, a transformation that requires approximately $189Np^4$ operations. Considerable effort has gone into developing fast translations schemes that require $O(p^2)$ operations. However, most of these schemes suffer from numerical instabilities.

COMBINE MULTIPOLE EXPANSIONS WITH PLANE WAVE EXPANSIONS The most recent algorithm for the three-dimensional case combines multipole expansions with exponential or plane wave expansions, which diagonalize the translation operators. Apparently this new method does not have numerical instabilities and requires $O(p^2)$ operations. Unfortunately, it introduces further mathematical complications related to the six plane-wave expansions (one for every face of the cube) and the expansion of the appropriate Green's function. With a convenient choice of the average number of particles contained in the leaf cells, the total operation count becomes $150Np + 5Np^2$. The reader is referred to Ref. (49) for further details.

BOUNDARY CONDITIONS The original FMM was developed for a single, nonperiodic cell. Schmidt & Lee (54) extended the FMM to periodic boundary conditions by combining it with Ewald techniques. At the end of the upward pass, all multipole expansions at all refinement levels of the unit cell are calculated. To implement the PBC, the unit cell is surrounded by image cells that, by definition, have the same multipole expansion around their centers. To proceed with the downward pass, one needs local expansions of the image cells. This is achieved through a multipole to local translation, which is implemented by an Ewald sum formulation. After this, the local expansions of all periodic image cells are available to evaluate the local expansion of the unit cell. The downward pass then proceeds as before. Lambert (55) derived an alternate approach, where the electrostatic potential and forces due to a finite periodic array of copies of the unit cell are evaluated by continuing the upward pass past the root cell to the successively larger set of copies. This can be considered a numerical implementation of the procedure of Smith (30).

Comparing the FMM with the Ewald Sum and Particle Mesh–Based Approaches

The relative efficiency of FMM versus Ewald summation and the various particle-mesh approaches has been the subject of some controversy. To our knowledge, the best study is that reported by Pollock & Glosli (19), because

they are the only group to attempt fully optimized implementations of each approach. They conclude that both P3M and FMM are more efficient than Ewald summation for macroscopic systems but that the P3M approach is significantly more efficient than FMM for any conceivable system size, despite the superior asymptotic scaling of the latter [$O(N)$ versus $O(N \log N)$]. They also discuss other advantages of the P^3M method, the most obvious one being the ease of coding. Also, noncubic periodic cells can be easily implemented with the particle-mesh methods, but not (easily) with the FMM. Finally, they point out that P3M can be used for nonperiodic systems by using a filter function in Fourier space, a fact that does not seem to be widely appreciated.

A few further comments are, however, warranted. First, the new methods introduced by Greengard & Rokhlin (49) may lead to substantial improvements in efficiency, although at the cost of even greater code complexity. Parallel implementations of FMM scale better than P3M or PME with the number of processors (J Board, personal communication). This is due to problems in efficiently parallelizing the FFT. The FMM may integrate better with multiple time-step methods (25), since the Ewald reciprocal sum involves neighboring high-frequency atom-atom interactions. The FMM approach is also better suited to problems involving highly nonuniform particle density because the tree approach lends itself well to an adaptive approach (57), whereas in these cases the grid size K in the P3M (see above) will grow faster than linearly with the number of particles. On the other hand, as noted above, FMM suffers from a lack of energy conservation unless very high accuracy is employed (22), whereas we will see below that PME has very good energy conservation properties. There are many other variants of the FMM, or of the particle-mesh methods, or combinations of both methods, that aim at improving efficiency. The reader is referred to Reference (39) for a review.

ARTIFACTS DUE TO EWALD SUMMATION

As discussed in the previous section, the use of Ewald summation and the particle mesh–based approximations has led to dramatic improvements in stability of biomolecular simulations. A concern voiced by many workers is that somehow the long-range periodicity artifactually stabilizes the system, i.e. that conformational fluctuations are inhibited by Ewald. Often, the lower root mean square (RMS) deviation from experimental structure seen in Ewald MD simulations, when compared to simulations using cutoffs, is accompanied by lower RMS fluctuations, although this is not always true. In a set of crystal simulations of the HIV-1 protease unit cell, York et al (58) compared the use of PME with that of short (9 Å) and long (18 Å) cutoffs. The short cutoff led to an unstable simulation, with the RMS deviation from the crystal structure increasing

monotonically with simulation time. The large cutoff led to a stable simulation, with RMS deviations only slightly larger than with PME. However, the RMS fluctuations were actually larger, and in better agreement with experimental B-factors, in the simulation using PME.

More recently, Smith & Pettitt (59, 60) have studied dynamical artifacts due to Ewald summation. They focused on the question of the size of barriers to free rotation of dipoles, due to long-range periodicity. Imagine a single ideal dipole at the origin of the unit cell, which is then replicated periodically, as discussed above. The potential energy of the system will depend on the dipole orientation, which is not appropriate for solution studies of biomoleculars. However, Smith & Pettitt showed that the rotational barriers are negligible for dipolar molecules in high dielectric solvents at room temperature, with typical simulation cell sizes. It will be interesting to see similar studies in low dielectric solvents.

As multiple-nanosecond simulations become routine, it will be possible to compare the time scale of simulated conformational transitions in peptides and flexible regions of proteins with those observed experimentally. A recent example of this type of study was given by Mohanty et al (61), who also employed the "locally enhanced sampling" method to accelerate conformational sampling of peptides. The results seen in these types of studies will depend on the force field, the integration algorithms used, including the methods used to control temperature and pressure, as well as possibly on the boundary conditions used. It will obviously be pointless to improve the agreement with experiment by tuning the force fields unless the other issues are understood, at least on the time scale under study. Schlick et al (72) have recently reviewed the various approaches to accelerating MD simulations by means of longer time steps. They point out a number of possible problems related to resonance phenomena. Artifacts in sampling due to subtle deficiencies in integration schemes may appear in long simulations.

Recently, Harvey et al (62) have pointed out severe artifacts due to temperature-control schemes in molecular dynamics. These can lead to loss of kinetic energy in certain degrees of freedom over the nanosecond time scale. Earlier, the phenomenon of hot solvent, cold protein was discovered and treated by means of separate temperature scaling. This phenomenon was later seen to be the result of cutoff artifacts, and separate temperature scales for solvent and solute are no longer necessary with Ewald summation techniques. However, an even more dramatic artifact emerged in an early implementation of the PME algorithm. The internal degrees of freedom cooled off monotonically, whereas the temperature scaling eventually caused the center of mass to acquire all of the kinetic energy. This gave origin to the "flying block of ice." In simulations of pure water, this is seen as a dramatic change in the diffusion coefficient as a function of time, which unfortunately is not apparent in short (subnanosecond) simulations. This artifact was traced to the incorrect treatment of attractive dispersion

interactions. When these are treated with cutoff through Verlet lists, it is important that the list be updated (effectively) every time step. This can be effected by using a list with a larger cutoff evaluating only those atom pairs within the real cutoff and updating the list when it is no longer valid due to atom movements.

In Figure 1, we show the results of a set of 5-nanosecond simulations of pure water with no temperature or pressure control (the NVE ensemble). The systems consisted of 216 SPC water molecules in cubic cells at standard density. The nonbonded cutoff was 8 Å without a switching function, electrostatics were handled by PME using cubic interpolation, a grid density of $\sim$1 Å and an Ewald coefficient of $\sim$0.35. The (leapfrog) Verlet time step was 2 femtoseconds, and an analytic version of SHAKE (63) was used. All simulations were performed using AMBER version 4.1 (64), modified to explore the effect of various list update strategies. The average temperature over 10-picosecond subintervals is

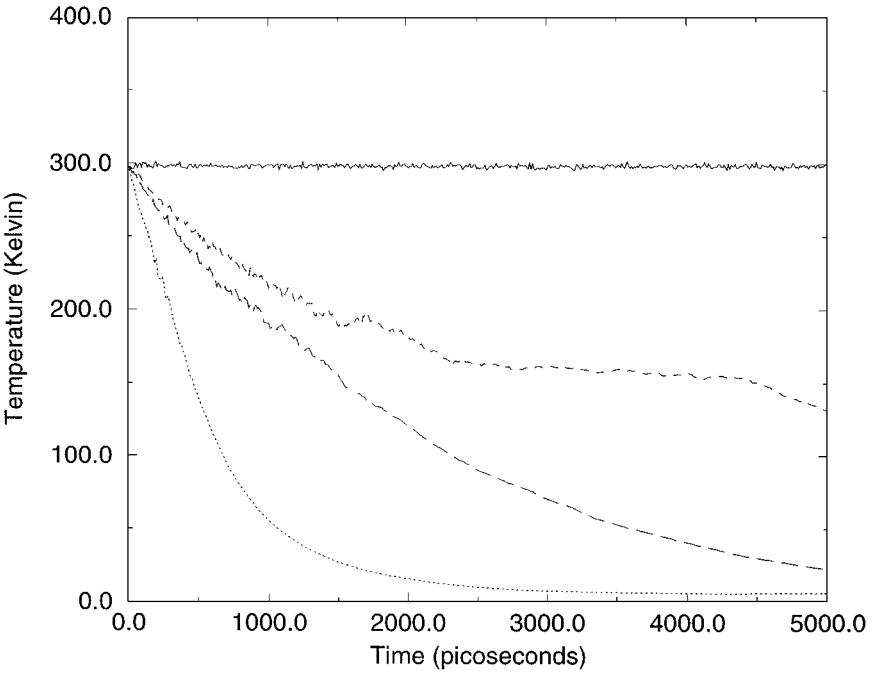

Figure 1 Effect of the nonbond list update time on system temperature in PME simulations of SPC water without temperature control (NVE). The running time average of temperature over 10 picosecond windows is plotted versus simulation time. The results are for update every step (*solid line*), every 5 steps (*dashed line*), every 10 steps (*long dashed line*), and every 20 steps (*dotted line*). The integration time step was 2 femtoseconds.

shown. Clearly, when the effective list was updated in every time step, there is no visible drift in the temperature over the 5 nanoseconds. This demonstrates that PME yields very good energy conservation. The total energy did, however, exhibit a small change of about 2 kcals mol^{-1} out of about -1750 kcals mol^{-1} over this time. When a one-femtosecond time step was used, the energy drift over 5 nanoseconds was reduced to about 0.6 kcals mol^{-1} (data not shown), in agreement with the expected quadratic dependence of the error of the Verlet integrator on time step. In addition, as expected, the diffusion constant, when calculated over 100-picosecond windows, exhibited no apparent drift over the course of the simulation (data not shown). In contrast, when the nonbonded list is not correctly updated, the energy decreases dramatically. These results do not reflect an artifact of Ewald summation or PME, since the results are even more dramatic when pure Lennard-Jones fluids are simulated (data not shown). It is easy now to see how the "flying block of ice" can arise if temperature control through velocity scaling is introduced into this system. These results also demonstrate the importance of energy conservation as a criterion of simulation quality. If temperature control is used to correct for lack of energy conservation, artifacts are almost certain to arise over long simulations.

Recently, several groups have studied the effect of boundary conditions on calculated free energies. A recent review is given by Levy & Gallicchio (65). Ionic charging free energies are particularly interesting. Due to the simplicity of the system, sampling limitations can be completely overcome, and agreement on force field parameters allow quantitative comparison between different approaches, unlike the more complex case of biomolecular simulations involving different force fields. By including the self energy ζ in Ewald calculations, Hummer et al (66) demonstrated remarkable system size consistency of ionic charging free energies. Note that the self energy need not be calculated in Monte-Carlo simulations; however, it is automatically included in typical implementations of Ewald summation in MD, such as with PME. Thus, in high dielectric solvents, standard Ewald summation results are only very weakly dependent on simulation cell size. Recently, Figuerido et al (56), Hummer et al (67), and Sakane et al (68) have derived finite size corrections for ionic charging in low dielectric solvents. These corrections are large, in contrast to the case in water.

Unfortunately, size consistency, while necessary, is not sufficient for a correct calculation. Size-consistent ionic charging free energies were reported earlier by Aqvist (69) using spherical cluster boundary conditions, with the SCAAS boundary potential (70). Recently, Darden et al (71) compared the charging free energies in water, obtained using Ewald (PME) in PBC, to those obtained in spherical boundary conditions. Size-consistent results were obtained in both cases, but the results differed substantially. For example, using

the Aqvist parameters for sodium-water interactions, a charging free energy of ~ -85 kcals mol^{-1} was obtained with Ewald summation, whereas with spherical boundary conditions, no cutoff of electrostatics, and inclusion of a Born correction a result of ~ -98 kcals mol^{-1} was obtained, close to the Aqvist result of ~ -100 kcals mol^{-1}. A priori, the spherical boundary conditions would seem more appropriate. However, the difference between the two sets of results was traced to the electric potential drop across the vacuum-water interface, an effect that is not accounted for by current boundary potentials. Using Gauss' law together with the spherical symmetry, a rigorous treatment of the charging process in cluster interiors is possible. The results using this treatment are in quantitative agreement with those using Ewald summation. This agreement was shown to hold for a variety of simple ions. Results using this treatment for ions in clusters of model low dielectric solvents were also in agreement with Ewald summation, if the above system size corrections were applied.

From the above study (71), we can infer that ionic charging free energies are not only size consistent, but correct, if Ewald summation is used. That is, the free energy is a correct function of the ion-water interaction potentials. These can now be fit appropriately to the experimental values. On the other hand, if spherical clusters are used, the surface potential of water must be accounted for. It will be interesting to see the effect of boundary conditions on solvation energies of more complex solutes.

PERSPECTIVE

With the continued growth of computer power and the advent of new algorithms for efficient treatment of long-range electrostatics, it is no longer necessary to resort to uncontrolled approximations such as the use of cutoffs. In the near future, the simulations using explicit solvents will most commonly be done using PBC with Ewald summation, some fast lattice summation or FMM approximation to Ewald summation, or perhaps some reaction-field treatment, where appropriate. Another popular choice will be the use of nonperiodic boundary conditions, such as spherical boundary conditions, with perhaps a reaction field to account for the missing bulk solvent beyond the sphere. The electrostatics for these isolated systems can be treated either by FMM, or as pointed out above, by the particle-mesh approach, using a filter in Fourier space.

Future computationally oriented studies will likely focus on the need for efficient parallel implementations of FMM and particle-mesh approaches. However, even more interesting work remains to be done to fully characterize the effects of long-range boundary conditions on biomolecular dynamics and energetics. As the study of force fields becomes more sophisticated, with the long-range goal of truly ab initio dynamics, the lessons learned about the

effect of boundary conditions will be of lasting value, since the long-range effects depend only on Coulomb's law.

Literature Cited

1. Berne BJ, Straub JE. 1997. Novel methods of sampling phase space in the simulation of biological systems. *Curr. Opin. Struct. Biol.* 7:181–89
2. Parrinello M. 1997. From silicon to RNA: the coming of age of ab initio molecular dynamics. *Solid State Commun.* 102:107–20
3. Briggs EL, Sullivan DJ, Bernholc J. 1996. A real-space multigrid-based approach to large-scale electronic structure calculations. *Phys. Rev. B* 54:14362
4. Leach AR. 1996. *Molecular Modeling, Principles and Applications*. Harlow, UK: Addison-Wesley Longman
5. Madden PA, Wilson M. 1996. 'Covalent' effects in 'ionic' systems. *Chem. Soc. Rev.* 25:339–50
6. van der Hoef MA, Madden PA. 1998. Novel simulation model for many-body multipole dispersion interactions. *Mol. Phys.* 94:417–33
7. Stone AJ. 1996. *The Theory of Intermolecular Forces*. Oxford, UK: Clarendon
8. Auffinger P, Westhof E. 1999. Molecular dynamics simulations of nucleic acids. In *Encyclopedia of Computational Chemistry*. New York: Wiley & Sons. In press
9. Cheatham TE III, Miller JL, Spector TI, Cieplak P, Kollman PA. 1997. Molecular dynamics simulations on nucleic acid systems using the Cornell et al force field and particle mesh Ewald electrostatics. In *ACS Symposium Series: Molecular Modeling and Structure Determination of Nucleic Acids*. Am. Chem. Soc.
10. Schreiber H, Steinhauser O. 1992. Cutoff size does strongly influence molecular dynamics results on solvated polypeptides. *Biochemistry* 31:5856–60
11. York DM, Yang W, Darden T, Pedersen LG. Towards the accurate modeling of DNA: the importance of long-range electrostatics. *J. Am. Chem. Soc.* 117:5001–2
12. Steinbach PJ, Brooks BR. 1994. New spherical cutoff methods for long-range forces in macromolecular simulation. *J. Comp. Chem.* 15:667–83
13. MacKerrell AD. 1997. Influence of magnesium ions on duplex DNA: structural, dynamic, and solvation properties. *J. Phys. Chem. B* 101:646–50
14. Feller SE, Pastor RW, Rojnuckarin A, Bogusz S, Brooks BR. 1996. Effect of electrostatic force truncation on interfacial and transport properties of water. *J. Phys. Chem.* 100:17011–20
15. Warshel A, Papazyan A. 1998. Electrostatic effects in macromolecules: fundamental concepts and practical modeling. *Curr. Opin. Struct. Biol.* 8:211–17
16. Smith W. 1982. Point multipoles in the Ewald summation. *CCP5 Inf. Q.* 4:13–25
17. Ewald P. 1921. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* 64:253–87
18. Hockney RW, Eastwood JW. 1981. *Computer Simulation Using Particles*. New York: McGraw-Hill
19. Pollock E, Glosli J. 1996. Comments on PPPM, FMM, and the Ewald method for large periodic Coulombic systems. *Comp. Phys. Comm.* 95:93–110
20. Darden TA, York DM, Pedersen LG. 1993. Particle mesh Ewald: an $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–92
21. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–93
22. Bishop T, Skeel R, Schulten K. 1997. Difficulties with multiple time stepping and fast multipole algorithm in molecular dynamics. *J. Comp. Chem.* 18:1785–91
23. Greengard L, Rokhlin V. 1987. A fast algorithm for particle simultations. *J. Comp. Phys.* 73:325–48
24. Board JA, Causey JW, Leathrum JF, Windemuth A, Schulten K. 1992. Accelerated molecular dynamics simulation with the parallel fast multipole algorithm. *Chem. Phys. Lett.* 198:89–94
25. Figuerido F, Levy R, Zhou R, Berne B. 1997. Large scale simulation of macromolecules in solution: combining the periodic fast multipole method with

multiple time step integrators. *J. Chem. Phys.* 106:9835–49
26. Allen MP, Tildesley DJ. 1987. *Computer Simulation of Liquids*. Oxford, UK: Clarendon
27. Kittel C. *Introduction to Solid State Physics*. New York: Wiley & Sons
28. Nose S, Klein M. 1983. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* 50:1055–76
29. DeLeeuw SW, Perram JW, Smith ER. 1980. Simulation of electrostatic systems in periodic boundary conditions I: lattice sums and dielectric constants. *Proc. R. Soc. Lond.* A373:27–56
30. Smith ER. 1994. Calculating the pressure in simulations using periodic boundary conditions. *J. Stat. Phys.* 77:449–72
31. Caillol JP. 1994. Comments on the numerical simulations of electrolytes in periodic boundary conditions. *J. Chem. Phys.* 101:6080–90
32. Hardy GH. 1975. *A Course of Pure Mathematics*. Cambridge, UK: Cambridge Univ. Press
33. Boresch S, Steinhauser O. 1997. Presumed versus real artifacts of the Ewald summation technique: the importance of dielectric boundary conditions. *Ber. Bunsenges. Phys. Chem.* 101:1019–29
34. Hammes-Schiffer S, Andersen H. 1994. A new formulation of the Hartree-Fock-Roothan method for structure calculations on crystals. *J. Chem. Phys.* 101:375–93
35. Perram JW, Petersen HG, DeLeeuw SW. 1988. An algorithm for the simulation of condensed matter that grows as the 3/2 power of the number of particles. *Mol. Phys.* 65:875–93
36. Natoli V, Ceperley DM. 1995. An optimized method for treating long-range potentials. *J. Comp. Phys.* 117:171–78
37. Procachi P, Marchi M. 1996. Taming the Ewald sum in molecular dynamics simulations of solvated proteins via a multiple time scale algorithm. *J. Chem. Phys.*
38. Tuckerman ME, Berne BJ, Martyna GJ. 1992. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* 97:1990–2001
39. Toukmaji A, Board JA. 1996. Ewald sum techniques in perspective: a survey. *Comp. Phys. Comm.* 95:78–92
40. Darden T, Toukmaji A, Pedersen L. 1997. Long-range electrostatic effects in biomolecular simulations. *J. Chem. Phys.* 94:1346–64
41. Deserno M, Holm C. 1998. How to mesh up Ewald sums: a theoretical and numerical comparison of various particle mesh routines. *J. Chem. Phys.* 109:7678–93
41a. Deserno M, Holm C. 1998. How to mesh up Ewald sums II: An accurate error estimate for the particle-particle particle-mesh algorithm. *J. Chem. Phys.* 109:7694–701
42. Ferrell R, Bertschinger E. 1994. Particle-mesh methods on the connection machine. *J. Mod. Phys.* 5:933–56
43. Hansen J. 1986. Molecular-dynamics simulations of Coulomb systems in two and three dimensions. In *Molecular Dynamics Simulation of Statistical-Mechanical Systems*, ed. G Ciccotti, WG Hoover. Amsterdam: North Holland
44. Greengard LF. 1988. *The Rapid Evaluation of Potential Fields in Particle Systems*. Cambridge, MA: MIT Press
45. Shimada J, Kaneko H, Takada T. 1993. Efficient calculations of Coulombic interactions in biomolecular simulations with periodic boundary conditions. *J. Comp. Chem.* 14:867–78
46. Luty BA, Davis ME, Tironi IG, van Gunsteren WF. 1994. A comparison of particle-particle particle-mesh and Ewald methods for calculating electrostatic interactions in periodic systems. *Mol. Simul.* 14:11–20
47. York D, Yang W. 1994. The Fast Fourier Poisson (FFP) method for calculating Ewald sums. *J. Chem. Phys.* 101:3298
48. Greengard LF. 1994. Fast algorithms for classical physics. *Science* 265:909–14
49. Greengard L, Rokhlin V. 1997. A new version of the fast multipole method for the Laplace equation in three dimensions. *Acta Numer.* 6:229–70
50. Jackson JD. 1975. *Classical Electrodynamics*. New York: Wiley & Sons
51. Appel AW. 1985. An efficient program for many-body simulations. *SIAM J. Sci. Stat. Comput.* 6:85–103
52. Barnes J, Hut P. 1986. A hierarchical $O(N \log N)$ force-calculation algorithm. *Nature* 324:446–49
53. Esselink K. 1995. A comparison of algorithms for long-range interactions. *Comp. Phys. Comm.* 87:375–95
54. Schmidt KE, Lee MA. 1991. Implementing the fast multipole method in three dimensions. *J. Stat. Phys.* 63:1223–35
55. Lambert CG, Darden TA, Board JA. 1996. A multipole-based method for efficient calculation of forces and potentials in macroscopic periodic assemblies of particles. *J. Comp. Phys.* 126:274–85
56. Figuerido F, Buono GD, Levy RM. 1997. On finite-size corrections to the free energy of ionic hydration. *J. Phys. Chem. B* 101:5622–23

57. Fenley M, Olson W, Chua K, Boschitsch A. 1994. Fast adaptive multipole method for computation of electrostatic energy in simulations of polyelectrolyte DNA. *J. Comp. Chem.* 17:976
58. York DM, Darden TA, Pedersen LG. 1993. The effect of long-range electrostatic interactions in simulations of macromolecular crystals: a comparison of the Ewald and truncated list methods. *J. Chem. Phys.* 99:8345–48
59. Smith PE, Pettitt BM. 1996. Ewald artifacts in liquid state molecular dynamics simulations. *J. Chem. Phys.* 105:4289–93
60. Smith PE, Pettitt BM. 1997. On the presence of rotational Ewald artifacts in the equilibrium and dynamical properties of a zwitterionic tetrapeptide in aqueous solution. *J. Phys. Chem. B* 101:3886–90
61. Mohanty D, Elber R, Thirumalai D, Beglov D, Roux B. 1997. Kinetics of peptide folding: computer simulations of STPFDV and peptide variants in water. *J. Mol. Biol.* 272:423–42
62. Harvey SH, Tan RK, Cheatham TE III. 1998. The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition. *J. Comp. Chem.* 19:726–40
63. Miyamoto S, Kollman P. 1992. Settle: an analytical version of the shake and rattle algorithms for rigid water molecules. *J. Comp. Chem.* 13:952–62
64. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE III, et al. 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Comm.* 91:1–41
65. Levy RM, Gallicchio E. 1998. Computer simulations with explicit solvent: recent progress in the thermodynamic decomposition of free energies, and in modeling electrostatic effects. *Annu. Rev. Phys. Chem.* 49:531–67
66. Hummer G, Pratt LR, Garcia AE. 1996. On the free energy of ionic hydration. *J. Phys. Chem.* 100:1206–15
67. Hummer G, Pratt LR, Garcia AE. 1997. Ion sizes and finite-size corrections for ionic-solvation free energies. *J. Chem. Phys.* 107:9275–77
68. Sakane S, Ashbaugh HS, Wood RH. 1998. Continuum corrections to the polarization and thermodynamic properties of Ewald sum simulations for ions and ion pairs at infinite dilution. *J. Phys. Chem. B* 102: 5673–82
69. Aqvist J. 1990. Ion-water interaction potentials derived from free energy perturbation simulations. *J. Phys. Chem.* 94:8021–24
70. King G, Warshel A. 1989. A surface constrained all-atom solvent model for effective simulations of polar solutions. *J. Chem. Phys.* 91:3647–61
71. Darden T, Pearlman D, Pedersen L. 1998. Ionic charging free energies: spherical versus periodic boundary conditions. *J. Chem. Phys.* 109:10921–35
72. Schlick T, Barth E, Mandziuk M. 1997. Biomolecular dynamics at long timesteps: bridging the timescale gap between simulation and experiment. *Annu. Rev. Biophys. Biomol. Struct.* 26:181–222

Annual Review of Biophysics and Biomolecular Structure
Volume 28, 1999

CONTENTS