Theor Chem Acc (1998) 99:279–288
DOI 10.1007/s002149800m13

Theoretical
Chemistry Accounts

Overview

Recent advances in molecular dynamics simulation towards the realistic representation of biomolecules in solution

Thomas E. Cheatham III, Bernard R. Brooks

Laboratory of Biophysical Chemistry, National Heart, Lung, and Blood Institute, 12A-2041, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892-5626, USA

Received: 9 April 1998 / Accepted: 21 May 1998 / Published online: 13 August 1998

Abstract. Coupled advances in empirical force fields and classical molecular dynamics simulation methodologies, combined with the availability of faster computers, has lead to significant progress towards accurately representing the structure and dynamics of biomolecular systems, such as proteins, nucleic acids, and lipids in their native environments. Thanks to these advances, simulation results are moving beyond merely evaluating force fields, displaying expected structural fluctuations, or demonstrating low root-mean-squared deviations from experimental structures and now provide believable structural insight into a variety of processes such as the stabilization of A-DNA in mixed water and ethanol solution or reversible β-peptide folding in methanol. The purpose of this overview is to take stock of these recent advances in biomolecular simulation and point out some common deficiencies exposed in longer simulations. The most significant methodological advances relate to the development of fast methods to properly treat long-range electrostatic interactions, and in this regard the fast Ewald methods are becoming the de facto standard.

Key words: Molecular dynamics – Biomolecular simulation

1 Introduction

Classical molecular dynamics (MD) simulation of biomolecules, such as proteins, nucleic acids, and lipids in their native environments, have seen significant advance due to the availability of reliable molecular mechanical (MM) force fields, development of better simulation methodologies and most notably reasonable treatment of the long-range electrostatic forces and faster computers. With these methods, the results move beyond simply testing force fields or displaying low root-mean-squared deviations (RMSd) from experimental structures to give believable structural and phenomenological insight into a variety of processes. Recent simulations suggest the methods are capable of representing subtle environmentally dependent conformational equilibria in DNA duplexes [1–4] and also provide an understanding of lipid structure at the atomic level which generalizes the results of experiments [5–8]. Moreover, these simulations aid in the prediction of peptide and protein-loop conformation and help interpret protein-folding mechanisms [9–14]. By moving beyond simply evaluating methods, the reasonable and reliable simulation of biomolecular systems gives a detailed picture of their structure, energetics, and dynamics which can ultimately provide insight into biological function. In addition to the routine use of restrained MD simulations as an aid in structural refinement of crystal or NMR structures, the dynamics can also reliably estimate anisotropic rotational diffusion tensors [15] and help in the interpretation of N-H order parameters [16], consistent with NMR data. Additionally these methods can give atomic-level insight into the short-time scale (ps-μs) dynamics typically masked by averaging in X-ray crystallographic or NMR spectroscopic experiments. Beyond directly complementing experiments, if the methods are proven to be reliable they can be applied in cases where experimentation is currently limited or difficult. Examples include the study of highly flexible systems such as membranes, or the investigation of ''prototype'' molecules not yet synthesized or characterized structurally (such as various backbone modifications to DNA that may be potentially useful as antisense therapeutics).

For the purposes of this overview, ''biomolecular simulation'' refers to investigation of the structure and dynamics of proteins, nucleic acids, lipids, or other biological molecules at an atomic level using theoretical techniques, such as MD or Monte Carlo methods [17–20], and presenting the most realistic representation of the environment as possible. The last point deserves elaboration. Since these molecules are typically large, and since accurate investigation further requires a detailed representation of the integral solvent and ionic environment, detailed treatments require a minimal

Correspondence to: T.E. Cheatham III

representation of the system and use of fairly simple empirical potential functions to keep the simulations tractable. The number of atoms in a standard simulation of a biomolecule (with explicit representation of the solvent) generally involves 1000 or more atoms (typically in the range of 10,000 to 50,000 atoms). This represents a moderate sized protein, small nucleic acid structure (< 100 base pairs) or minimal lipid bilayer (< 100 lipids) surrounded by $\sim$5–15 Å of solvent. It should be remembered that this is effectively a minimal system and as such may have solvent properties that differ considerably from bulk, effective concentrations that are rather high (in periodic boundary simulations) and may contain an insufficient number of molecules to represent certain large-scale phenomena, such as lipid-bilayer undulations [21]. Moreover, given that the typical representation includes only a single biomolecule, certain processes such as aggregation of DNA at high ethanol concentrations cannot easily be studied. For systems of this size, MD simulations in the 1–50 nanosecond range (using reliable and accurate methods) currently represent the state-of-the-art. This time range is reasonable for equilibration of the initial solute and solvent structure (which normally occurs in the 50 ps to multi-nanosecond range) and limited sampling of thermally accessible conformations. Empirical (pairwise) MM potential functions are necessary due to the large number of energy and force evaluations required since small integration time steps are typically applied to properly represent the high-frequency motions and allow stable integration (for a good review, see Schlick et al. [22]). Integration steps in the 1–2 fs range are routine, although more recently various groups are starting to apply slightly longer time steps (in the 2–6 fs range) through the application of multiple time-step methods [23] such as RESPA [24], Langevin-Newton, Implicit-Euler [22], or split integration symplectic methods [25]. Due to their stability, reversibility, and reasonable representation of the Hamiltonian, RESPA methods are the most promising to increase the speed of these calculations. However, without artificially increasing the masses to limit the high-frequency motions, time steps longer than $\sim$5 fs may be unfeasible when explicit water is included in the calculation because of the anharmonicity of the diffusive modes, stiff interactions of the van der Waals potential, water libration motion at $\sim$750 cm^{-1}, and integration resonance phenomena. This 5 fs time step limitation manifests itself as a lack of reasonable energy conservation and typically represents half the value of the Verlet integration catastrophe time step (or the period of the highest frequency motion divided by π) [22, 26].

The size of the systems investigated, inclusion of explicit solvent (necessary on some level due to its structural role), and the large number of energy and force evaluations required place the demands of these simulations well beyond the capabilities of ab initio methods at present. And even with the emerging fast parallelized semi-empirical methods [27, 28] nanosecond-length MD will not be possible in the near future. However, despite the simplified nature of the empirical pairwise potentials, force fields are actually quite sufficient for reasonable representation of the structure and dynamics. Although the use of an empirical potential precludes the use of these methods to investigate chemical processes, such as enzyme mechanism, in these cases it is possible to use hybrid quantum mechanical (QM) and MM treatments (QM/MM) where only a limited core part of the system is treated quantum mechanically [29–31]. However, the extreme computational demands of these calculations has limited extensive study and there are still a number of open issues in QM/MM treatments as how best to merge the QM and MM regions, what parameters to use for the empirical potential, and what level of treatment to use in the QM region (such as specially parameterized semi-empirical treatments, ab initio, etc.).

2 Strict energy conservation, stable integrators and modern pressure treatments

The few detailed QM/MM investigations have emphasized issues important to any MD treatment, specifically the need for strict energy conservation, stable, reversible, and ideally symplectic integrators, and reasonable treatments of pressure and temperature. The various forms of the Verlet algorithm, such as the leap-frog, have all these properties and are the most commonly employed integrators. Despite having these properties, minor changes to the precise details of the integration scheme can have a profound effect on the derived properties. For example, with the Verlet algorithms, the computation of the temperature can be performed with either the on-step or half-step velocities. Whereas the temperature calculated with the on-step velocities is independent of the vibrational frequencies, the commonly used half-step velocities can significantly overestimate the temperature leading to reduced "real" temperatures in the simulation when temperature coupling is used. Similarly with pressure, it is important to use the half-step velocities when calculating the viral in order to maximize the cancellation of errors [32]. To better maintain the temperature and pressure, the simulation community has begun to move away from the "standard" weak-coupling method [33] for temperature and pressure coupling to more advanced treatments, such as Nosé-Hoover thermostated chains [10, 24] or Langevin piston methods [34]. These methods not only allow for better temperature and pressure coupling in the correct ensemble [35], but also allow for new ensembles which may be more appropriate for the atomic systems under study, such as constant surface tension or surface area ensembles appropriate for the simulation of lipid bilayers [36].

Stability is important in longer simulations otherwise the gradual accumulation of errors from minor deficiencies in the methods can lead to artifactual behavior. An example of this relates to the use of methods which uniformly scale the velocities to maintain temperature coupled with incomplete energy conservation when the center of mass motion is not removed during the dynamics. If potential energy is lost, the coupling to the temperature bath will slowly scale the velocities up to maintain the desired temperature. This scales up the center of mass motion. In the case of a periodic system,

the center of mass translational energy cannot couple back into the system. This not only leads to a violation of equipartition, but the center of mass translation will continue to grow until essentially all of the molecular motion is within this degree of freedom or any low-frequency mode that does not couple well (such as methyl group rotation) [37]; with a very small energy drain, the system (be it a periodic box of water or more complicated system involving a solvated biomolecule) resembles a "flying block of ice" by ~500–1000 ps. In general, if a method is employed that results in an energy drain that is spatially located (i.e., group-based and the phenomena of truncation with water crossing back and forth across the cutoff which leads to the "hot water/cold protein" problem) or related to a specific vibrational frequency (i.e., the weak-coupling constant pressure algorithm which results in an energy drain from the highest frequency motion), and uniform scaling is employed to restore lost energy, then poorly coupled low-frequency modes can acquire a significant amount of excess energy which can drastically alter observables. A useful technique to identify these problems is to calculate atomic temperature and fluctuations (which requires memory but little extra computation) with some post-simulation analysis. It was not until longer simulations (nanosecond length) were routine and methods to accurately treat the long-range electrostatics were routinely applied (which eliminated the more serious cutoff induced heating) that these effects were noticed in periodic boundary simulations. With commonly applied methods, small energy drains are routinely possible:

1 If buffered pairlists and conservative or heuristic updates are not applied to prevent the omission of atoms entering or leaving the cutoff sphere.
2 SHAKE tolerances are not stringent enough.
3 The weak-coupling method is used for pressure control.
4 Time steps are too large.
5 Approximations are used to represent the long-ranged electrostatic interactions.

In the latter case, it is important not only to conserve energy but conserve forces and to this end it may be more appropriate to apply fast Ewald methods which interpolate the electrostatic forces rather than the energy [38]. Therefore, care should also be taken when applying any of the new fast methods to represent long-range electrostatic interactions (which are discussed briefly in a later section). For example, accuracy may be lost if low-order multipole expansions are applied with fast multipole methods (FMM) [39] or if the direct and reciprocal interactions are unbalanced in Ewald simulations or inaccurate charge grids applied with the fast Ewald methods. A simple means to eliminate the problem related to transfer of energy to center of mass motion (rotation or translation) or low-frequency modes is to remove or repartition this excess energy. While this is straightforward for center of mass degrees of freedom (which should only be removed as appropriate for the system under consideration, such as only removing center of mass translation in periodic systems), it is not straightforward for low-frequency modes. Although removal of the center of mass motion in periodic systems seems to eliminate the major artifactual behavior, drains of energy from selected modes, coupled with uniform energy scaling to pump energy back into the system, leads to inhomogeneity and could lead to artifacts in longer simulations. Therefore, care should be taken to conserve energy. Even with conservation of energy however, uniform temperature scaling to maintain temperature can lead to a frequency dependence on temperature scaling [37] with the weak-coupling approach. One approach to avoid this problem is to alter the usual calculation of the scaling factor (within a Verlet leap-frog integration scheme), λ, as follows:

$$\lambda = \left[1 + C_v \left(\frac{k_B}{2} \right) \frac{\Delta t}{\tau} \left(\frac{2k_B T_{\text{ref}} - \langle \sum m v_n^2 \rangle}{\left\langle \sum \frac{m}{2} \left(v_{n+\frac{1}{2}}^2 + v_{n-\frac{1}{2}}^2 \right) \right\rangle} \right) \right]^{1/2}$$

where T_{ref} is the reference temperature, v_n is the on-step velocity and $v_{n-1/2}$, $v_{n+1/2}$ are the previous and current half-step velocities, m represents the masses in a sum over all atoms, Δt the time step, τ the coupling time, C_v the heat capacity and k_B is the Boltzmann constant. Use of this scaling avoids the frequency dependence and can be used for any size system (including a harmonic oscillator) in contrast to the standard method which does not couple properly for small system sizes.

Accurately calculating pressure is also extremely important since small errors can also accumulate. Unfortunately, this is not straightforward since the methods used to calculate the pressure, typically based on the internal virial and calculated as the dot product of the coordinates with the internal forces [17, 20], are fraught with pitfalls. With this in mind, the pressure calculated is best tested on systems with known answers, such as a small system in a large box (which should have zero pressure). Care also should be taken to be compatible with the integrator, such as using half-step velocities in a Verlet scheme, and it should be noted that any addition of forces to the system, such as random and frictional forces of Langevin dynamics, can have a profound effect on the virial. Another example where difficulties arise is when restraints or constraints are applied to fix part of the system. When applying harmonic positional restraints or fixing particular coordinates, the calculated force does not represent an internal force and therefore is not included in the calculated (internal) pressure. Typically this leads to an overestimation of pressure end can lead to serious artifacts upon change in the box size (due to pressure coupling). In longer runs with part of the system fixed or restrained to initial coordinates (generally performed for the purposes of equilibrating the system), the overestimation of pressure can lead to artificial expansion of the periodic box leading to the appearance of vacuum "bubbles". The solution is not to use these types of restraints/constraints in constant pressure calculations or to convert these into internal forces. The latter can be accomplished by applying relative harmonic positional restraints (i.e., the reference coordinates are best fit to the current coordinates at each time step and the energy and now internal forces to these

relative coordinates calculated, including the force from the RMS fitting which can be derived from the law of cosines applied in 3N dimensions). Another issue with constant pressure calculations is how to estimate properties such as diffusion which will be influenced by the coordinate scaling during pressure coupling. A simple means to eliminate this non-physical motion is to post-process the trajectory to scale the individual coordinate frames back to the average box size; thus removing the effect of box-size fluctuations. Failure to do this results in an overestimation of the motion of particles near the edge of the periodic box.

Of course instability and inaccuracy are only a major issue if the errors are systematic, as in the above examples. Random force errors, such as those resulting from the slow accumulation of errors due to limited numerical precision of the computer, do not seem to lead to artifacts since these errors are equally likely to add as subtract from the total energy. To this end, differences obtained between parallel and sequential MD runs resulting from differences in the order of operations do not lead to significant errors and simply manifest the inherently chaotic nature of the integration as is discussed in detail by Braxenthaler et al. [40].

3 The need for proper representation of the environment

Ideally, one would like to use the best possible representation in the simulation and this represents a trade-off in computational cost and accuracy. Since water is an integral part of the structure of biological macromolecules, some representation is clearly necessary. Without any representation of the solvent, or even with simple distance-dependent dielectric treatments which attempt to mimic solvent screening, DNA structure tends to distort [41] and proteins compact [42]. Less expensive than explicitly including solvent, implicit water models can give reasonable energetic insight, such as the estimation of the pKa of titratable groups, solvation and binding free energies, and salt effects on nucleic acid structure [43]. However the omission of structural water and difficulty in accurately representing the hydrophobic effect limits the utility. In spite of this, there has been some limited use of continuum methods, such as Poisson Boltzmann methods, in MD simulation [44, 45]. Despite the additional cost, more reliable representations of the structure and dynamics are seen when explicit solvent is included in the simulation. However, this raises issues about how to include the solvent, what type of boundary conditions to apply, and what solvent model to use.

The boundary conditions applied can be broken into two classes based on whether the system is non-periodic, or in other words surrounded by a finite amount of explicit solvent with a vacuum, continuum or other implicit model beyond this, or periodic where the finite system is effectively replicated in a periodic lattice and toroidal boundary conditions are applied [17]. The point of discussing the nature of the boundary in this overview is that there are issues and potential for artifacts with both types of boundary condition. In non-periodic systems, the most serious artifacts occur when the system is surrounded by a vacuum interface and when methods are not applied to break up the water ordering at the surface. This surface tension leads to artificially large pressures at the center of the system and reduced fluctuations [46, 47]. A crude estimation of the internal pressure increase is $R\Delta P \approx 15{,}000$ Å-atm in a droplet of water with a radius of R Å. A commonly applied method to reduce the severity of this is the application of stochastic boundary conditions where random forces and/or simple boundary forces are applied to water near the surface [48, 49]. However, removal of this ordering at the surface is non-trivial and potentially incomplete with stochastic methods. This has led to the development of various complicated, water model and droplet-size-dependent, functional forms applied to treat the molecules at the boundary that have been used with limited success [50]. An alternative means to avoid the ordering at the surface and include some representation of the missing solvent polarization is to eliminate the vacuum interface, by surrounding the explicit system with an implicit solvent model, such as an array of Langevin dipoles, a dielectric continuum or a reaction field [51–54]. Issues with these simulations include avoiding force discontinuities at the interface between the explicit and implicit systems and the computational cost of providing a realistic representation. Despite the availability of these methods, they have seem limited application in large-scale biomolecular simulation. Part of the reason for this is that often within the solvent blob all of the pairwise interactions are represented (to avoid cutoff artifacts). This becomes very expensive, even more expensive than periodic boundary conditions (PBC), for solvated biomolecules which tend to be rather large. Of course, it is possible to apply fast multipole algorithms [55] or even non-periodic fast Ewald methods [56] to make these calculations more tractable and competitive with PBC. In a recent paper the solvent boundary potential of Beglov & Roux was applied with a cell-multiple treatment [57]. However, the biggest problem is that none of these boundary potentials seem to completely remove the effect of the interface.

PBC seem to be the preferred boundary conditions, particularly thanks to the recent availability of fast methods to calculate the long-range interactions (as is discussed in the next section). With PBC, the calculations are typically facilitated by calculating pairwise interactions only to atoms within a given distance or cutoff. Two types of cutoffs are commonly employed, spherical cutoff (which can be performed on an atom to atom, charge group or residue basis) and minimum image conventions. With minimum image cutoffs, all interactions of a given atom with other atoms within the periodic unit (centered on the given atom) are included, whereas with a spherical cutoff, only those atoms/groups within a given radius are included. The minimum image methods have been shown to have serious artifacts (due to the presence of more interactions in the corners of the periodic unit) in the simulation of dipolar solutions, such as significantly anisotropic and damped reorientational motion [58]. These type of boundary conditions are best avoided. With spherical cutoffs, the largest problem relates to truncation of the long-range interactions, most

notably the electrostatics interactions. As discussed in more detail elsewhere [46] and in the next section, large artifacts are readily apparent depending on the way the interactions are truncated. These artifacts range in severity from complete disruption of the structure (such as is the case in the simulation of nucleic acids with truncated group-based cutoffs) to completely inhibited motion (observed with narrow-ranged atom-based switching of the potential). As in the case of non-periodic boundaries, it is possible to apply a reaction-field treatment for interactions outside the cutoff sphere in periodic boundary simulations [59–62]. However, the approaches based on extensions to the Onsager theory [63] break down if electroneutrality is not maintained within each possible cutoff sphere (due to the ill-defined dipole) and any treatment which assumes a uniform dielectric outside the cutoff sphere will not be appropriate in cases where individual biomolecules in the system are bigger than the cutoff sphere (as is often the case in large-scale biomolecular simulation). Additionally with group-based methods, molecules entering and leaving the cutoff sphere or effectively moving into and out of the continuum represent a discontinuity in the forces which leads to lack of energy conservation. Although some of these deficiencies can be minimized by using larger cutoff spheres, the computational cost then becomes prohibitive with respect to the fast truly periodic methods. Therefore, these methods have not seen considerable use in large-scale biomolecular simulation.

An alternative to applying a cutoff to the long-range electrostatic interactions is to apply a truly periodic method, such as Ewald summation [17, 64]. However, the imposition of true periodicity can lead to artifacts since all the atoms now fully interact with their periodic images. In principle, this may lead to correlation of fluctuations, such as the inhibition of the free rotation of a dipole (as would be expected in the absence of true periodicity) and artifactual forces between ions in a periodic box [65]. For example, two ions separated by half the box length in a periodic box will experience no net attractive or repulsive force due to balancing interactions with the periodically imaged ions. In principle, this artifact can be eliminated through artificial construction of the simulation cell and the use of sine fast Fourier transforms which modulate the effect of the true periodicity [66]. However, this comes at the expense of other observables, such as conservation of energy, and the procedure has seen limited application. Alternatively, the periodicity can be effectively removed in Ewald simulations by placing a finite system (which means the same issues regarding an explicit interface as discussed above apply) in a larger periodic box and applying an appropriate filter function [56], although this technique has yet to see use in biomolecular simulation. Fortunately, the artifacts from true periodicity seem to be minor. In solvents with a sufficient permittivity, such as water, dipolar rotation is not strongly inhibited. This has been shown in a series of simulations by Smith and Pettitt where the effective difference between free rotation in solution and the hindered rotation expected with true periodicity is less than kT for model dipoles in water and moreover the rotational diffusion of a small zwitterionic peptide in water is close to what is expected [67–69]. Similarly, the potential of mean force as a function of separation for two ions in water does not display force artifacts in boxes as small as $\sim$12 Å [70]. Further evidence comes from estimation of the conformational PMF of a blocked trialanine peptide in a 26 Å periodic box (in vacuo) which agrees reasonably well to that obtained without a cutoff and no periodicity [10]. Additionally, reasonable size-independent free energies of solvation can be calculated, even for charged systems [71–73]. These results suggest that the artifacts of true periodicity are likely to be small for net-neutral periodic cells when solvent with a sufficiently high dielectric is used.

Another issue with truly periodic methods relates to the treatment of the dielectric boundary, a term analogous to a reaction field that comes out of the solution of the Poisson equation under PBC (the Ewald potential). This term depends on the surface, shape, and composition of the periodic system at the limits of the summation over the macroscopic system (not to be confused with the surface of the unit cell) [74]. In many implementations, *tin-foil* or conducting boundary conditions are assumed which effectively represent a dielectric boundary with a dielectric constant of infinity and thereby set this reaction field term to zero. Simulations by Boresh and Steinhauser show that imposition of *tin-foil* boundary conditions may lead to overstabilization of the correlation between dipoles at larger distances [75]. They therefore recommend surrounding the system by a dielectric boundary with a dielectric closer to that of the experimental system. However, given that the surface structure and therefore the polarization effects should vanish for any disordered macroscopic system, it has been argued that *tin-foil* boundary conditions are appropriate for liquid simulations, whereas non-conducting boundary conditions should be applied in crystal simulations [58].

A final issue worth mentioning is that many force fields include specific terms (such as van der Waal parameters or solvent parameters) that are parameterized using cutoffs or for use in simulations applying cutoffs (i.e., the commonly applied TIP3P [76] and SPC/E [77] water models). A concern is that complete treatment of the long-range interactions could in principle lead to artifacts from the force field. However, typically much better behavior is seen in Ewald simulations, such as consistently closer agreement to experimental structures when Ewald treatments are applied [47].

4 The importance of properly treating long-range electrostatics

Inclusion of the long-range electrostatic interactions is critical for poly-ionic systems such as lipids and nucleic acids. This has been well known in biomolecular simulation ever since the first MD on nucleic acids where it was necessary to eliminate the phosphate charges to prevent distortion of the DNA duplex in short simulations [78]. Similar behavior is observed today even with state-of-the-art force fields when poor

cutoff methods, such as residue-based truncation in the 8–20 Å range, are applied. In this case and without artificial restraints added to maintain Watson-Crick base pairing, distortion and disruption of duplex DNA is observed within ~ 200 ps of MD [79]. By minimizing the electrostatic force discontinuities at the cutoff (through the use of atom-based force shifts at the cutoff or other methods) stable nanosecond-length simulation of DNA duplexes [80] in solution is possible, however it has been observed that computed transport properties are very sensitive to long-range electrostatic cutoffs. For lipid simulation [5] even a 16 Å atom-based force shifted cutoff provided significantly altered transport properties relative to an Ewald treatment. Given that a reasonable representation of the structure of these systems is possible, and moreover the fact that reasonable simulation of protein systems has been noted for many years despite the application of questionable cutoff methods, one might question the importance of including long-range electrostatic interactions. However, given the fact that long-range interactions can now be included at little or no additional cost (as discussed below), the proper "physics" is represented, and true periodicity artifacts appear to be small, we question why not use Ewald methods? In fact, this seems to be the general consensus with more and more research groups routinely applying the fast Ewald methods due to their generality and availability in many of the common MD codes, such as CHARMM [81], AMBER [82], and GROMOS [83].

The major reason for continued use of standard cutoff methods is in part because of their legacy and in part because of the greater ease of generating very scaleable parallel implementations and ease of use for simple modeling projects. Moreover, the problems with cutoff simulations have been largely masked because of the short-time scale of the simulations, fortuitous cancellation of errors, and in the case of proteins because the ionic interactions are mostly confined to the surface residues which do not strongly influence the structure. This fortuitous cancellation of errors in cutoff simulations compared to Ewald simulations has been seen even in the simulation of charged ions in solution where some properties, such as orientational correlation functions and some transport properties are in reasonably good agreement [58]. Better examples of this fortuitous cancellation of errors are seen in the simulation of an α-helical peptide by Schreiber & Steinhauser [84, 85]. Although reasonable representation of the α-helix was observed with a commonly applied 10 Å cutoff (or with an Ewald treatment), disruption of the helical structure was seen at shorter (6 Å) or longer (14 Å) cutoffs. This shows that the effect of the cutoff is not monotonically related to cutoff distance and therefore better results are not necessarily achieved by applying longer cutoffs. Of course, the fortuitous cancellation of errors is not always observed and in many cases cutoff simulations routinely display artifactual behavior due to the omission of the long-range electrostatics. This includes incorrect long-range orientational correlations, strong anticorrelation of dipolar fluctuations [86] and decreased translational and rotational motion [60]. A particularly graphic example of a cutoff-induced artifact, even when a spline-smoothed potential at the cutoff (11.8 Å) is applied, was seen in simulations by Bader and Chandler where a net-attractive PMF between two like-charged ions in solution was seen [70]. In these simulations, a well at ~6–7 Å and an attractive potential beyond 7 Å are observed in cutoff simulations, in contrast to Ewald methods which possess no well or attractive PMF (despite the imposition of true periodicity). The attractive potential is strong enough that during dynamics the charged ions separated by ~9 Å move closer together in the cutoff simulations. This type of behavior was also observed independently by Dang and Pettitt for chloride ion pairing in solution when a cutoff was applied [87] which is not observed in corresponding Ewald simulations [88]. In addition to force artifacts between charged ions, in some conditions the transport properties are also drastically altered in cutoff simulations. This has been seen in the simulation of lipid bilayers, where long-range order was apparent in the radial distribution functions at long distances, the electrostatic profile across the membrane was grossly misrepresented, and greater viscosity and lower translational diffusion of the water observed [5]. Although most of these difficulties were well known, particularly by the liquid-state simulation community who has been using Ewald methods for years, these have been largely ignored by the biomolecular simulation community due to the large amount of solvent that needs to be included in the calculations and the masking of errors, as discussed above. This rationalization has largely changed thanks to the availability of faster computers and fast Ewald methods.

The first large-scale application of fast Ewald methods in large-scale nanosecond-length biomolecular simulation was presented by Darden and co-workers who investigated a variety of nucleic acid and protein crystals using the efficient particle mesh Ewald (PME) method [38, 89]. Darden developed the method to overcome the artifactual gross distortion of DNA structure seen when residue-based cutoffs were applied. With these methods, excellent representation of the structure of crystals of nucleic acids and proteins were observed [90–94]. The PME method is a generalization of the particle-particle mesh Ewald (PPPM) method inspired by Hockney and Eastwood [95] and developed more recently by Luty et al. [96]. An additional fast Ewald method is the fast Fourier-Poisson method [93]. All of these methods have in common the use of fast Fourier transforms to speed the calculation of the reciprocal space interactions. The PME method has become the most widely applied of the fast Ewald methods due to its accuracy, speed, and facility for constant pressure calculations [38, 97]. A good comparison of PME and PPPM methods has been recently published by Darden and co-workers [98]. With the fast Ewald methods, effectively shorter cutoffs in the 9–11 Å range can be applied to build the pairlist (since the cutoff effectively only applies to the Lennard-Jones interactions) which makes these methods faster than the standard atom-based force-shifted cutoffs employed in the 12–14 Å range. Given that the artifacts apparent in the simulation resulting from the true periodicity appear less severe than those in comparable cutoff simulations, fast Ewald methods should be routinely applied. Alter-

natives to the fast Ewald methods, which can also be used in truly periodic simulations, are periodic FMM [55, 99–102]. Despite the apparent formally better scaling (linear in number of atoms, N, rather than the N-log(N) scaling seen with fast Ewald methods), these methods realistically scale as N-log(N) to maintain consistent force errors as the system size increases and have not seen as much usage in accurate nanosecond-length biomolecular simulation to date. Part of this stems from the complexity of the code and need for high-order multipole expansions to obtain sufficient accuracy [26] and also since the break-even point for the computational cost occurs at larger numbers of atoms than are typically included in current biomolecular simulations. For simulations in the $\sim 10^4$ atom range, the fast Ewald methods are typically faster (assuming an equivalent level of accuracy) [101].

5 Empirical pairwise potential functions: How can such as simple model accurately represent the structure and dynamics of biomolecules?

The accuracy of the potential function largely relates to the quality and specifics of the parameterization. Increases in computer power significantly aid the parameterization efforts, not only by allowing higher level basis sets and representation of larger molecules in QM treatments of model compounds, but by allowing more in-depth evaluation of representative biomolecules in MD and this in turn leads to better force fields. While the arguably more strongly parameterized force fields such as MM3 [103] or the Merck Molecular force field [104] perform better for strained and diverse small molecules, the "work horse" force fields for biomolecular simulation are still the specifically parameterized force fields such as Cornell et al. [105] (proteins and nucleic acids), CHARMM/MacKerell [106] (proteins, nucleic acids, and lipids), OPLS [107, 108] (organic liquids, carbohydrates) and others. Part of the success of the more recent force fields is that they provide a balanced treatment of not only the intramolecular interactions, but are parameterized to provide proper balance with the intermolecular interactions, such as those with solvent. Moreover, the application of methods which properly represent the long-range electrostatic forces has also tremendously increased the reliability of the simulations. A nice example is seen in simulations of ubiquitin, where the results with the Cornell et al. [105] force field are consistently better than the earlier Weiner et al. [109] force field (in terms of RMSd to the crystal structure) and in both cases genuinely improve (i.e., the improvement is not due to the structure stiffening) when Ewald methods are applied [47]. Although each force field has specific weaknesses, progress is underway to improve the deficiencies. With further methodological advances (to increase the speed of calculations with explicit polarization), it is likely that force fields for biomolecular systems which include explicit polarization will emerge in the next few years. Of course, success of the methods comes down to comparison with experimental values and ultimately prediction, and significant progress towards this goal has been seen.

6 Realistic biomolecular simulation with an explicit representation of the environment

To properly represent the structure and dynamics of biomolecular systems, it is desirable to include some representation of the environment in the simulations. Since water and counterions often play a structural role, some explicit representation is necessary and when an explicit solvent is included, the simulations can reasonably represent the structure and dynamics. The availability of the fast Ewald methods has lead to a renaissance in the simulation of highly charged systems, such as nucleic acids. Previously, the simulation of nucleic acids was plagued by instabilities. Now routine and reliable nanosecond-length simulation of various nucleic acid structures is possible. A non-exhaustive list of successes has involved investigating ion association [80, 110], sequence-specific structure and hydration [1, 111, 112], reliability of the force fields [113–116], various backbone modifications (such as PNA [117], phosphoramidate [118], and guanidine [119] modified nucleic acids) and photodamage [120, 121] in nucleic acid duplexes. Other structures, such as DNA triplexes [122–125], RNA hairpin loops [126, 127] and t-RNA [128, 129] have also been investigated with reasonable success.

Highlights of the simulation of DNA duplexes are the demonstration that environmentally dependent conformational preferences can be reasonably represented. Whereas spontaneous A-DNA to B-DNA transitions are observed in water with the Cornell et al. force field [1], simulations with the same DNA duplex and force field in $\sim$ 85% ethanol show preferential stabilization of A-DNA [2, 4] and spontaneous B-DNA to A-DNA transitions are seen when 4:1 hexaammine cobalt (III) is added to particular sequences in agreement with experimental data [3]. These simulations suggest that water and ion association in the major groove is likely to stabilize A-DNA and provide molecular-level insight that has not been easily seen in experimental A-DNA solution studies. What is particularly exciting about these calculations is the ability to sample relevant nucleic acid conformations in MD simulations and preferential stabilization based only on subtle changes to the environment.

Similar levels of success have also been seen in protein and lipid-bilayer simulation. Simulations are now able to critically assess the effective surface area per head group and reasonably represent the structure and interfacial properties of bilayers of various types [5–8]. Most exciting in the representation of protein dynamics is the demonstration of the ability rather than simply to unfold a protein, to move closer to the correct structure from misfolded structures. This has been seen with small peptide models [9] and also more recently with small proteins in solution. In the latter case, instabilities due to cutoff effects lead to unfolding of the protein whereas correct behavior is observed in Ewald simulations (Carlos Simmerling, personal communication). In addition, exciting work has continued in the investigation of reaction coordinates for protein folding [11, 12] and even more recently with reversible-folding simulations of

small β-peptides in methanol [13, 14]. These latter simulations show correct folding to the native structure and in high-temperature simulations allow characterization of the unfolded states. These simulations suggest that there is only a relatively small number of "unfolded" conformations (in contrast to the expected exponential explosion) which suggests that given sufficient computer time (and sampling) understanding of the folding process may be possible.

Major issues that still compound the simulation, beyond the short-time and length scales, relate to conformational sampling problems and the difficulty in overcoming any but rather small barriers to structural transition in nanosecond-length MD. For example, MD results suggest that imaginary structures (i.e., metastable structures which have never been observed experimentally) such as B-RNA are stable over a multi-nanosecond simulation [113]. Methods are needed to achieve effective lowering of the conformational barriers or to allow structures across the barriers in an unbiased manner. A promising technique that effectively lowers the barriers is the locally enhanced sampling methodology [130]. This allows spontaneous transformation of incorrect to correct RNA hairpin-loop structures [131], reasonable representation of slightly misfolded structures, and prediction of protein-loop conformation (as discussed previously).

7 Commodity processors, powerful workstations and massively parallel supercomputers: Biomolecular simulation for the masses?

Although perhaps less glamorous than the methodological and force-field improvements, the steady increase in computer power has greatly facilitated these advances has allowed more detailed investigation of the potentials and methods, allowed sampling over longer time scales, and highlighted previously masked deficiencies as previously discussed. Generous grants of computer time and access to massively parallel computers at the various supercomputer centers throughout the U.S. and Europe have aided the field. The availability of these machines has led to widespread availability of parallelized MD codes and moreover most of the recent work discussed herein could not have been performed without access to this kind of supercomputing power. With these large machines, nanosecond-length simulations can currently be performed in hours or days rather than months or years. A benefit of the fast-moving pace of the computer industry is that simulations which previously required access to the fastest available supercomputers a few years ago can now be performed on relatively inexpensive workstations. This brings realistic biomolecular simulation into the realm of more general users, allowing anyone with a computer and a reasonable understanding of the methods to investigate the dynamics of biomolecular systems in a fairly realistic fashion. These increases in computer power allow investigation on a level beyond simple molecular modeling, allowing some limited conformational sampling and inclusion of accurate methods and potentials. The microprocessor revolution, coupled with the low cost of commodity processors, further aids this effort [132]. The relatively modest communication requirements of standard MD algorithms allows efficient "Beowulf" class computers [133], or loosely coupled networks of commodity processors, to be built on a modest budget which provide reasonable computational power for biomolecular simulation. An example of such a computer is the lots of boxes on shelves (LoBoS) computer system being developed at the NIH out of Pentium Pro processors and a matched 100 baseT ring topology network used primarily for biomolecular simulation with CHARMM, AMBER, and GAMESS (see http://www.lobos.nih.gov). These modest costs bring a machine of this type into the cost level affordable by individual researchers. This is not to say that more computer power is not necessary and further access to the state-of-the-art parallel computer systems warranted since ideally we would like to push biomolecular simulation into the microsecond to millisecond time scale with more realistic energy representations and the study of more than minimally sized biomolecules. Access to greater computational power will bring molecular simulation into the time-scale range appropriate for studying more complex processes such as protein folding, to properly represent lateral diffusion in lipid bilayers [134], or to study more complicated conformational transitions in nucleic acids such as the B-DNA to Z-DNA transition. It is hoped that continued access will be made available to these state-of-the-art computational resources to bring structural biomolecular simulation up to the status of more traditional experimental tools.

Acknowledgements. We thank Drs. van Gunsteren, Langley, and Simmerling for making preprints of their work available. We also thank Eric Billings, Tom Darden, Peter Kollman, David Langley, Rich Pastor, Carlos Simmerling, and Pete Steinbach for many helpful discussions and Carolina Reyes for a critical reading of this manuscript.

References

1. Cheatham TE III, Kollman PA (1996) J Mol Biol 259:434
2. Cheatham TE III, Crowley MF, Fox T, Kollman PA (1997) Proc Natl Acad Sci USA 94:9626
3. Cheatham TE III, Kollman PA (1997) Structure 5:1297
4. Langley DR (1998) J Biomol Struct Dyn (submitted)
5. Feller SE, Pastor RW, Rojnuckarin A, Bogusz S, Brooks BR (1996) J Phys Chem 100:17011
6. Feller SE, Yin D, Pastor RW, MacKerell ADJ (1997) Biophys J 73:2269
7. Feller SE, Venable RM, Pastor RW (1997) Langmuir 13:6555
8. Tobias DJ, Tu K, Klein ML (1997) Curr Opin Colloid Interface Sci 2:115
9. Mohanty D, Elber R, Thirumalai D, Beglov D, Roux B (1997) J Mol Biol 272:423
10. Samuelson S, Tobias DJ, Martyna GJ (1997) J Phys Chem B 101:7592
11. Guo Z, Brooks CL III, Boczko EM (1997) Proc Natl Acad Sci USA 94:10161
12. Sheinerman FB, Brooks CL III (1998) Proc Natl Acad Sci USA 95:1562
13. Daura X, van Gunsteren WF, Rigo D, Jaun B, Seebach D (1997) Chem Eur J 3:1410

14. Daura X, Jaun B, Seebach D, van Gunsteren WF, Mark AE (1998) J Mol Biol (submitted)
15. Tjandra N, Feller SE, Pastor RW, Bax A (1995) J Am Chem Soc 117:12562
16. Chatfield DC, Szabo A, Brooks BR (1998) J Am Chem Soc 120, 5301–5311
17. Allen MP, Tildesley DJ (1987) Computer simulation of liquids. Oxford University Press, Oxford
18. Brooks CLI, Karplus M, Pettitt BM (1988) In: Prigogine I, Rice SA (eds) Proteins. A theoretical perspective of dynamics, structure, and thermodynamics, vol. 71. Wiley, New York, p 259
19. Harvey SC (1989) Proteins 5:78
20. van Gunsteren WF, Berendsen HJC (1990) Angew Chem Int Ed Engl 29:992
21. Feller SE, Pastor RW (1997) In: Pacific symposium on biocomputing '97. Altman RB, Dunker AR, Hunter L, Klein TE (eds). World Scientific, Maui, Hawaii, p 142–150
22. Schlick T, Barth E, Mandziuk M (1997) Ann Rev Biophys Biomol Struct 26:181
23. Streett WB, Tildesley DJ, Saville G (1978) Mol Phys 35:639
24. Tuckerman M, Berne BJ, Martyna GJ (1992) J Chem Phys 97:1990
25. Janezic D, Merzel F (1997) J Chem Inf Comput Sci 37:1048
26. Bishop TC, Skeel RD, Schulten K (1997) J Comp Chem 18:1785
27. Dixon SL, Merz KMJ (1997) J Chem Phys 107:879
28. Lee TS, York DM, Yang WT (1996) J Chem Phys 105:2744
29. Eurenius KR, Chatfield DC, Brooks BR, Hodoscek M (1996) Int J Quantum Chem 60:1189
30. Stanton RV, Little LR, Merz KM (1995) J Phys Chem 99:17344
31. Liu HY, Mullerplathe F, van Gunsteren WF (1996) J Mol Biol 261:454
32. Pastor RW, Brooks BR, Szabo A (1988) Mol Phys 65:1409
33. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) J Comp Phys 81:3684
34. Feller SE, Zhang Y, Pastor W, Brooks BR (1995) J Chem Phys 103:4613
35. Hoover WG (1985) Phys Rev A 31:1695
36. Zhang Y, Feller SE, Brooks BR, Pastor RW (1995) J Chem Phys 103:10252
37. Harvey SC, Tan RKZ, Cheatham TE III (1998) J Comp Chem (in press)
38. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) J Chem Phys 103:8577
39. Ding HQ, Karasawa N, Goddard WA (1992) J Chem Phys 97:4309
40. Braxenthaler M, Unger R, Auerbach D, Given JA, Moult J (1997) Proteins 29:417
41. Beveridge DL (1993) In: Westhof E Water and biological molecules. Beveridge DL, Swaminathan S, Ravishanker G, Withka JM, Srinivasan J, Prevost C, Lovise-May S, Longley DR, D. Capus FM, Botton PH (eds), Macmillan p. 165
42. Steinbach PJ, Brooks BR (1993) Proc Natl Acad Sci USA 90:9135
43. Honig B, Nicholls A (1995) Science 268:1144
44. Gilson MK, McCammon JA, Madura JD (1995) J Comp Chem 16:1081
45. Smart JL, Marrone TJ, McCammon JA (1997) J Comp Chem 18:1750
46. Steinbach PJ, Brooks BR (1994) J Comp Chem 15:667
47. Fox T, Kollman PA (1996) Proteins 25:315
48. Berkowitz ML, McCammon JA (1982) Chem Phys Lett 90:215
49. Brooks CLI, Brunger A, Karplus M (1985) Biopolymers 24:843
50. Essex JW, Jorgensen WL (1995) J Comp Chem 16:951
51. Warshel A, Levitt M (1976) J Mol Biol 103:227
52. King G, Warshel A (1989) J Chem Phys 91:3647
53. Papazyan A, Warshel A (1997) J Phys Chem B 101:11254
54. Beglov D, Roux B (1994) J Chem Phys 100:9050
55. Greengard L, Rokhlin V (1989) Chemica Scripta 29A:139
56. Pollock EL, Glosli J (1996) Comp Phys Comm 95:93
57. Sugita Y, Kitao A (1998) Proteins 30:388
58. Roberts JE, Schnitker J (1995) J Phys Chem 99:1322
59. Schreiber H, Steinhauser O (1992) J Mol Biol 228:909
60. van Gunsteren WF, Berendsen HJC, Rullman JAC (1978) Faraday Disc Chem Soc 66:58
61. Tironi IG, Sperb R, Smith PE, van Gunsteren WF (1995) J Chem Phys 102:5451
62. Tironi IG, Luty BA, van Gunsteren WF (1997) J Chem Phys 106:6068
63. Onsager L (1936) J Am Chem Soc 58:1486
64. Ewald P (1921) Ann Phys (Leipzig) 64:253
65. Valleau JP, Whittington SG (1977) In: Berne BJ (ed) Statistical mechanics, A. A modern Theoretical chemistry. Plenum, New York, p 137
66. Luty BA, van Gunsteren WF (1996) J Phys Chem 100:2581
67. Smith PE, Pettitt BM (1991) J Chem Phys 95:8430
68. Smith PE, Pettitt BM (1996) J Chem Phys 105:4289
69. Smith PE, Blatt HD, Pettitt BM (1997) J Phys Chem B 101:3886
70. Bader JS, Chandler D (1992) J Chem Phys 96:6423
71. Hummer G, Pratt LR, Garcia A (1997) J Chem Phys 107:9275
72. Figueirido F, Del Buono GS, Levy RM (1997) J Chem Phys 101:5622
73. Bogusz S, Cheatham TE III, Brooks BR (1998) J Chem Phys 108:7070–7084
74. DeLeeuw SW, Perram JM, Smith ER (1980) Proc R Soc London A373:27
75. Boresch S, Steinhauser O (1997) Bunsenges Phys Chem 101:1019
76. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926
77. Berendsen HJC, Grigera JR, Straatsma TP (1987) J Phys Chem 91:6269
78. Levitt M (1983) Cold Spring Harbor Symp Quant Biol 47:251
79. Cheatham TE III, Miller JL, Fox T, Darden TA, Kollman PA (1995) J Am Chem Soc 117:4193
80. MacKerell AD Jr. (1997) J Phys Chem B101:646
81. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) J Comp Chem 4:187
82. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, Ferguson D, Seibel G, Kollman P (1995) Comp Phys Comm 91:1
83. van Gunsteren WF, Berendsen HJC (1987) Groningen molecular simulation (GROMOS) library manual, BIOMOS: Nijenborgh, Groningen, The Netherlands
84. Schreiber H, Steinhauser O (1992) Biochemistry 31:5856
85. Schreiber H, Steinhauser O (1992) Chem Phys 168:75
86. Neumann M (1983) Mol Phys 50:841
87. Dang LX, Pettitt BM (1987) J Am Chem Soc 109:5531
88. Hummer G, Soumpasis DM, Neumann M (1993) Mol Phys 81:1155
89. Darden TA, York DM, Pedersen LG (1993) J Chem Phys 98:10089
90. Lee H, Darden TA, Pedersen LG (1995) J Chem Phys 102:3830
91. Lee H, Darden TA, Pedersen LG (1995) Chem Phys Lett 243:229
92. York DM, Yang W, Lee H, Darden TA, Pedersen L (1995) J Am Chem Soc 117:5001
93. York D, Yang W (1994) J Chem Phys 101:3298
94. York DM, Wlodawer A, Pedersen LG, Darden TA (1994) Proc Natl Acad Sci USA 91:8715
95. Hockney RW, Eastwood JW (1981) Computer simulation using particles. McGraw-Hill, New York
96. Luty BA, Tironi IG, van Gunsteren WF (1995) J Chem Phys 103:3014
97. Petersen HG (1995) J Chem Phys 103:3668
98. Darden TA, Toukmaji A, Pedersen LG (1997) J Chem Phys 94:1346

99. Greengard L (1994) Science 265:909
100. Figueirido F, Levy RM, Zhou RH, Berne BJ (1997) J Chem Phys 106:9835
101. Challacombe M, White C, Head-Gordon M (1997) J Chem Phys 107:10131
102. Schmidt KE, Lee MA (1997) J Stat Phys 89:411
103. Allinger NL, Yuh YH, Lii JH (1989) J Am Chem Soc 111:8551
104. Halgren TA (1996) J Comp Chem 17:490
105. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr., Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117:5179
106. Mackerell AD, Wiorkiewiczkuczera J, Karplus M (1995) J Am Chem Soc 117:11946
107. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) J Am Chem Soc 118:11225
108. Damm W, Frontera A, Tirado-Rives J, Jorgensen WL (1997) J Comp Chem 18:1955
109. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona GS, Profeta J, Weiner P (1984) J Am Chem Soc 106:765
110. Young MA, Jayaram B, Beveridge DL (1997) J Am Chem Soc 119:59
111. Pastor N, Pardo L, Weinstein H (1997) Biophys J 73:640
112. Duan Y, Wilkosz P, Crowley M, Rosenberg JM (1997) J Mol Biol 272:553
113. Cheatham TE III, Kollman PA (1997) J Am Chem Soc 119:4805
114. Young MA, Ravishanker G, Beveridge DL (1997) Biophys J 73:2313
115. Feig M, Pettitt BM (1997) J Phys Chem 101:7361
116. Cheatham TE III, Kollman PA (1998) In: Sarma M, Sarma R (eds) Structure, motion, interactions and expression of biological macromolecules. Adenine Press, Schenectady p 99
117. Sen S, Nilsson L (1998) J Am Chem Soc 120:619
118. Cieplak P, Cheatham TE III, Kollman PA (1997) J Am Chem Soc 119:6722
119. Luo J, Bruice TC (1997) J Am Chem Soc 119:6693
120. Miaskiewicz K, Miller J, Cooney M, Ossman R (1996) J Am Chem Soc 118:9156
121. Spector TI, Cheatham TE III, Kollman PA (1997) J Am Chem Soc 119:7095
122. Luo J, Bruice TC (1998) J Am Chem Soc 120:1115
123. Weerasinghe S, Smith PE, Pettitt BM (1995) Biochemistry 34:16269
124. Weerasinghe S, Smith PE, Mohan V, Cheng YK, Pettitt BM (1995) J Am Chem Soc 117:2147
125. Shields GC, Laughton CA, Orozco M (1997) J Am Chem Soc 119:7463
126. Miller JL, Kollman PA (1997) Biophys J 73:2702
127. Miller JL, Kollman PA (1997) J Mol Biol 270:436
128. Auffinger P, Westhof E (1997) J Mol Biol 269:326
129. Auffinger P, Louise-May S, Westhof E (1996) J Am Chem Soc 118:1181
130. Roitberg A, Elber R (1991) J Chem Phys 95:9277
131. Simmerling C, Miller JL, Kollman PA (1998) J Am Chem Soc (submitted)
132. Tirado-Rives J, Jorgensen WL (1996) J Comp Chem 17:1385
133. Sterling T (1996) Commun ACM 39:11
134. Pastor RW, Feller SE (1996) In: Merz KMJ, Ronx B (eds) Biological membranes. A molecular perspective from computation and experiment. Birkhauser, Boston, p 3