

The effect of longrange electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods

Darrin M. York, Tom A. Darden, and Lee G. Pedersen

Citation: *The Journal of Chemical Physics* **99**, 8345 (1993); doi: 10.1063/1.465608

View online: <http://dx.doi.org/10.1063/1.465608>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jcp/99/10?ver=pdfcov>

Published by the [AIP Publishing](#)

Articles you may be interested in

[A fast multipole method combined with a reaction field for long-range electrostatics in molecular dynamics simulations: The effects of truncation on the properties of water](#)

J. Chem. Phys. **118**, 10847 (2003); 10.1063/1.1574774

[Effects of long-range electrostatic potential truncation on the free energy of ionic hydration](#)

J. Chem. Phys. **106**, 8135 (1997); 10.1063/1.473800

[Molecular dynamics free energy simulations: Influence of the truncation of longrange nonbonded electrostatic interactions on free energy calculations of polar molecules](#)

J. Chem. Phys. **101**, 7953 (1994); 10.1063/1.468222

[A local reaction field method for fast evaluation of longrange electrostatic interactions in molecular simulations](#)

J. Chem. Phys. **97**, 3100 (1992); 10.1063/1.462997

[The extended Ewald method: A general treatment of longrange electrostatic interactions in microscopic simulations](#)

J. Chem. Phys. **89**, 3751 (1988); 10.1063/1.454897



AIP | Applied Physics
Letters

is pleased to announce **Reuben Collins**
as its new Editor-in-Chief



The effect of long-range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods

Darrin M. York^{a)}

University of North Carolina, Chapel Hill, North Carolina 27599-3290 and National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

Tom A. Darden

National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

Lee G. Pedersen

University of North Carolina, Chapel Hill, North Carolina 27599-3290 and National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

(Received 22 June 1993; accepted 14 September 1993)

Simulations of the HIV-1 protease unit cell using a 9 Å cutoff, 9/18 Å “twin-range” cutoff, and full Ewald sums have been carried out to 300 ps. The results indicate that long-range electrostatic interactions are essential for proper representation of the HIV-1 protease crystal structure. The 9 Å simulation did not converge in 300 ps. Inclusion of a 9/18 Å “twin-range” cutoff showed significant improvement. Simulation using the Ewald summation convention gave the best overall agreement with x-ray crystallographic data, and showed the least internal differences in the time average structures of the asymmetric units. The Ewald simulation represents an efficient implementation of the Particle Mesh Ewald method [Darden *et al.*, *J. Chem. Phys.* **98**, 10 089 (1993)], and illustrates the importance of including long-range electrostatic forces in large macromolecular systems.

I. INTRODUCTION

A major approximation made in conventional molecular dynamics simulations involves the neglect of long-range electrostatic forces.¹⁻³ Calculation of the Coulombic interaction between all minimum image pairs of atoms in the system is an N^2 computational problem, where N is the total number of atoms. The “minimum image convention” thus becomes restrictive for large macromolecular systems (typically $N > 10\,000$ atoms). An approximation which simplifies this problem involves truncating the Coulombic potential (r^{-1}) at a fixed cutoff distance, R_c , so that the sum over all pairs is reduced to a sum over a list of nonbond atoms (the Verlet list). This results in an order N procedure. Typically, values of R_c are between 8–10 Å. Increasing the cutoff to include longer range interactions quickly becomes restrictive since the computational requirement is proportional to R_c^3 . An alternative technique is to employ a “twin-range” cutoff.⁴ The “twin-range” method requires specification of a short- and a long-range cutoff, R_c^{short} and R_c^{long} . At the time of the nonbond list update, the Coulombic force at each atom i , \mathbf{f}_i , is calculated to the long-range cutoff, R_c^{long} , and separated into short- and long-range components, $\mathbf{f}_i = \mathbf{f}_i^{\text{short}} + \mathbf{f}_i^{\text{long}}$. At subsequent time steps, the short range component, $\mathbf{f}_i^{\text{short}}$, is recomputed from the nonbond list, and the long-range component, $\mathbf{f}_i^{\text{long}}$, remains fixed until the next nonbond list update when it is recomputed. The “twin-range” method assumes high frequency components of the long-range Coulombic forces are negligible.

An alternative to the truncation methods is to calculate the full electrostatic energy of the unit cell in a macroscopic lattice of repeating images using the Ewald sum-

mation convention.⁵ The total electrostatic energy, E^{el} , of a neutral unit cell U , containing N point charges q_1, q_2, \dots, q_N , located at positions $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$, can be written:^{6,7}

$$E^{\text{el}} = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N q_i q_j [\psi_E(\mathbf{r}_{ij}) + f(\mathbf{r}_{ij})] + J(\mathbf{D}, P, \epsilon') \quad (1)$$

where the functions $\psi_E(\mathbf{r})$ and $f(\mathbf{r})$ are defined by

$$\psi_E(\mathbf{r}) = \sum_{|\mathbf{n}| \neq 0} \frac{\text{erfc}(\beta|\mathbf{r} + \mathbf{n}|)}{|\mathbf{r} + \mathbf{n}|} + \frac{1}{\pi V} \sum_{|\mathbf{m}| \neq 0} \frac{\exp(-\pi^2 m^2 / \beta^2)}{m^2} \exp(2\pi i \mathbf{m} \cdot \mathbf{r}) \quad (2)$$

and

$$f(\mathbf{r}) = \begin{cases} \text{erfc}(\beta|\mathbf{r}|)/|\mathbf{r}|, & \mathbf{r} \neq 0 \\ -2\beta/\sqrt{\pi}, & \mathbf{r} = 0 \end{cases} \quad (3)$$

where $\text{erfc}(x)$ is the complementary error function, V is the volume of the unit cell U , and β is a positive parameter. The value of β adjusts the relative rates of convergence of the direct space and reciprocal space sums in Eq. (2), but is otherwise arbitrary. The second term $J(\mathbf{D}, P, \epsilon')$ in Eq. (1) was derived by de Leeuw *et al.*⁶ and depends on the total dipole moment \mathbf{D} of the unit cell, the shape P of the macroscopic crystal lattice, and the dielectric constant ϵ' of the surrounding medium. The infinite sums in Eq. (2) are over direct space (\mathbf{n}) and reciprocal space (\mathbf{m}) lattices, respectively.

TABLE I. Parameters of the HIV-1 PR crystal simulation.

Space group	$P4_12_12$
(a, b, c) (Å)	(50.24, 50.24, 106.56)
α, β, γ (°)	(90, 90, 90)
Density	1.17 g/cm ³
No. atoms	29 661
Grid dim. (K_1, K_2, K_3) ^a	(64, 64, 128)
Approx. grid size ^a	0.84 Å
No. recip. vectors ^a	102 347
β^a	0.386 Å ⁻¹

^aParameters used in the Ewald simulation using the PME method (Ref. 8).

In the past, simulation of macromolecular crystals using the Ewald expression for the total electrostatic energy was not feasible due to the large computational cost of evaluating Eq. (1). Recently a fast $N \times \log(N)$ algorithm for computing Ewald sums has been introduced by Darden *et al.*⁸ The present work presents the first application of the Particle Mesh Ewald (PME) method to a simulation of a large macromolecular crystalline system.

We have performed molecular dynamics (MD) simulations of the unit cell of HIV-1 protease (HIV-1 PR) for 300 ps using a 9 Å cutoff, a 9/18 Å “twin-range” cutoff, and the Ewald summation convention. Structures and fluctuations predicted from the simulations are compared with each other and with the 2.8 Å x-ray crystallographic structure of Wlodawer *et al.*⁹ The HIV-1 protease represents a well-studied biologically significant protein that is particularly well suited for testing the effects of inclusion of long-range electrostatics. HIV-1 PR is a 99 amino acid monomer which contains no stabilizing disulfide bonds. The protein is only moderately ionic, hence one might expect discrepancies between simulations to be amplified in systems of more highly charged proteins and particularly DNA. Furthermore, the HIV-1 PR unit cell contains a large percentage of solvent (~60%), and hence results reported here may have relevance to simulations in solution as well.

II. METHODS

Simulation methodology. Molecular mechanics and dynamics calculations were performed using AMBER3.0 Rev. A¹⁰ modified to incorporate the twin-range force correction,¹¹ and Ewald sums using the Particle Mesh Ewald method.⁸ The all-atom force field¹² was employed for amino acid residues. Solvent was treated explicitly using the rigid TIP3P model.¹³ A total of 16 chloride counterions were included to neutralize the net +2 charge on each monomer.¹⁴ Hydrogen bond lengths were constrained using a modified version of the SHAKE algorithm suggested by Ryckaert *et al.*¹⁵ to allow 1 fs integration time step. Counterion placement, water packing, equilibration, and energy refinement of the unit cell to arrive at the starting configuration have been detailed previously.¹⁴ Initial conditions were identical in each of the MD simulations. Parameters used for calculating the Ewald energies with the PME method are given in Table I. The surrounding dielectric was taken to be infinite (“tin-foil” boundary condi-

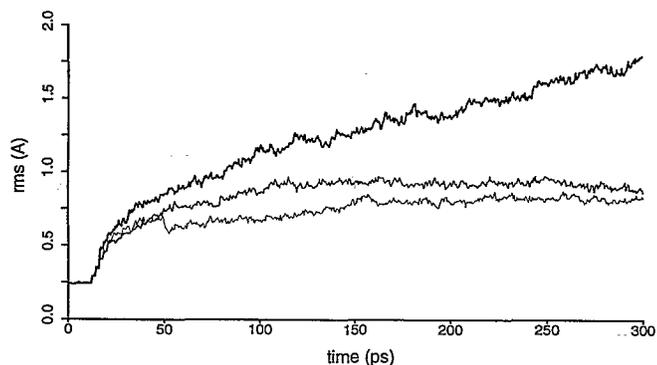


FIG. 1. Time evolution of the rms backbone deviation from the crystallographic structure of the instantaneous unit cell average structures for the 9 Å simulation (thick line), 9/18 Å “twin-range” simulation (medium line), and Ewald simulation (thin line).

tions) making the surface term in Eq. (1) vanish. van der Waals interactions were calculated using a 9 Å atom-based nonbond list in the Ewald simulation, and a 9 Å residue-based list in the cutoff simulations. All simulations were performed in the NVT ensemble at the crystallographic temperature (293 K), and carried out to 300 ps.

Unit cell. Simulations were based on the crystallographic structure of the HIV-1 protease reported by Wlodawer *et al.*⁹ Unit cell parameters and parameters used in computing the Ewald sum with the PME method are given in Table I. The unit cell contains 8 asymmetric units, each asymmetric unit consists of a single protein monomer. The experimental crystal density was 1.17 g/cm³ (assumed to be that of the isomorphous crystal reported by Navia *et al.*¹⁶), corresponding to approximately 60% solvent. This solvent concentration is on the high end of the range normally observed for protein crystals.¹⁷

III. RESULTS

Structural equilibration. Figure 1 shows the time evolution of the root-mean-square (rms) backbone deviation of the MD structures from the crystallographic structure. At each time point, an instantaneous unit cell average was calculated by transforming the coordinates of each monomer to a local principal axis system and averaging over all 8 monomers. The 9/18 Å and Ewald simulations equilibrate by about 150 ps with rms values of approximately 0.9 and 0.8 Å, respectively. The rms values for the 9 Å simulation did not equilibrate, but increased linearly from 75–300 ps, with a maximum value of 1.8 Å at 300 ps. The striking difference between the simulations clearly demonstrates that longer range electrostatic interactions are necessary for structural stability in protein crystals. Similar effects have also been observed in solution simulations of large proteins,¹⁴ and small peptides.³

Average structures. Average structures were calculated from 200–300 ps for the 9/18 Å and Ewald simulations. Time averages of individual monomers as well as unit cell averages (averages over all asymmetric units and time) were calculated. Since the 9 Å simulation did not equilibrate, subsequent discussion of this simulation is omitted.

The rms backbone deviation of the unit cell average structures from the crystallographic structure was 0.9 Å for the 9/18 Å simulation and 0.8 Å for the Ewald simulation. Although the rms values for the unit cell average structures are similar, examination of the rms deviations of the individual monomers reveals interesting differences. The average rms backbone deviation of the monomers from the crystallographic structure was 1.32 ± 0.31 Å for the 9/18 Å simulation and 1.13 ± 0.11 Å for the Ewald simulation. The average intermonomer rms deviation (deviation of the monomers with other monomers in the unit cell) was 1.51 ± 0.26 Å for the 9/18 Å simulation and 1.21 ± 0.15 Å for the Ewald simulation. Hence, the time average monomers of the 9/18 Å simulation are more structurally divergent than those of the Ewald simulation; i.e., the structures are less similar (as measured by the backbone rms) to each other and to the crystallographic structure than the corresponding structures in the Ewald simulation. In both simulations the individual monomers are more similar to the crystallographic structure than to one another. This suggests the monomers occupy localized regions of configuration space with corresponding average structures that bracket the crystallographic structure. Consequently, the unit cell average structures are considerably closer to the crystallographic structure than the individual time average monomers.

Secondary structure. Analysis of the secondary structure of the unit cell average structures was performed using the Kabsch and Sander program DSSP.¹⁸ Of the 62 secondary structural assignments made for the crystallographic monomer, 52 (84%) were preserved in both the 9/18 Å and Ewald unit cell average structures. The largest discrepancy between secondary structural assignments in the crystallographic structure and the unit cell average structures occurred at the dimer interface formed by the amino- and carboxyl-terminal residues. These residues make up a four stranded antiparallel beta sheet. Although hydrogen bonding at the dimer interface is well preserved in the simulations, this region shows a high degree of thermal motion, and the backbone configuration of the resulting unit cell averages was not assigned beta structure. If these regions are neglected (residues 1–4, 95–99), the over-

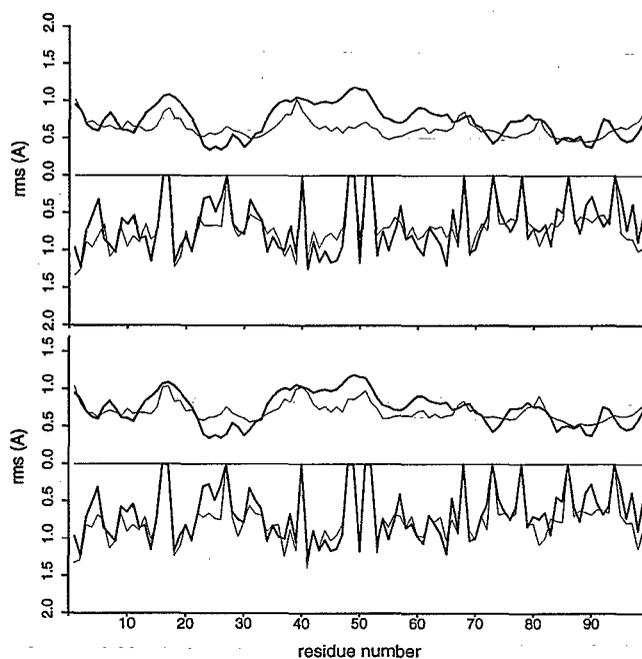


FIG. 2. Atomic fluctuations (rms) estimated from the isotropic temperature factors (thick lines) and calculated from the MD (thin lines) for the 9/18 Å simulation (top) and Ewald simulation (bottom). Residue averages are shown for backbone heavy atoms (above the rms=0 axis) and side chain heavy atoms (below the rms=0 axis).

all secondary structural homology increases to 93% for both unit cell averages.

Thermal fluctuations. Atomic fluctuations result from the thermal motion of atoms which vibrate in a local minimum potential of mean force. Experimentally, atomic fluctuations can be estimated from the isotropic temperature factors by the relationship $[\langle r_i^2 \rangle - \langle r_i \rangle^2]^{1/2} = [(3/8\pi^2) B_i]^{1/2}$, where r_i is the displacement vector of atom i , and B_i is the corresponding crystallographic B value. Figure 2 shows the residue averaged atomic fluctuations estimated from the isotropic temperature factors and calculated from the MD simulations for backbone atoms (above the abscissa) and side chain atoms (below the abscissa). The average backbone fluctuation calculated for the 9/18 Å simulation is slightly smaller than for the

TABLE II. Energy statistics for the average electrostatic energy $\langle E_{\text{ecl}} \rangle$, total energy $\langle E_{\text{tot}} \rangle$, temperature $\langle T \rangle$, and drift in total energy per step $\langle E_{\text{drift}} \rangle$.^a

Simulation	$\langle E_{\text{ecl}} \rangle$	$\langle E_{\text{tot}} \rangle$	$\langle T \rangle$	$\langle E_{\text{drift}} \rangle$	CPU ^b
9 Å	-1.33×10^5 (567)	-6.90×10^4 (861)	304.3(5.3)	2.27(1.70)	6.80 ^c
	-1.34×10^5 (390)	-6.91×10^4 (742)	303.9(5.3)	2.78(1.43)	7.68 ^d
9/18 Å	-1.33×10^5 (694)	-6.92×10^4 (1080)	306.4(6.5)	3.88(1.32)	12.50 ^c
	-1.34×10^5 (148)	-7.06×10^4 (220)	297.7(2.0)	0.56(1.57)	29.64 ^d
Ewald	-1.32×10^5 (121)	-6.87×10^4 (18.4)	293.2(1.1)	-0.06(0.004)	15.92 ^c
	-1.32×10^5 (130)	-6.87×10^4 (8.9)	293.4(1.1)	-0.03(0.004)	16.43 ^d

^aConstant energy MD runs (1 ps) restarted at 300 ps; energies are in kcal/mol; rms values are in parentheses.

^bParent processor CPU time for a dedicated 8 processor SGI-380; limited memory on this machine resulted in decreased throughput in the case of the Ewald simulation.

^c20 step nonbond list update.

^d5 step nonbond list update.

Ewald simulation (0.64 Å and 0.70 Å, respectively). Fluctuations of side chain atoms were slightly higher than for the backbone atoms in each case (0.79 Å and 0.86 Å, respectively). The absolute accuracy of fluctuations estimated from the crystallographic data is difficult to assess due to additional contributions to the Debye–Waller factors such as lattice disorder. However, the relative precision of these estimates has been shown to be meaningful.¹⁷ One can measure the degree of correlation between the experimentally derived fluctuations “*x*” and fluctuations calculated from the MD “*y*” by calculating the linear correlation coefficient $r_c = \langle \Delta x \Delta y \rangle / \langle \Delta x \rangle^{1/2} \langle \Delta y \rangle^{1/2}$, where $\Delta x = x - \langle x \rangle$. The correlation coefficient has values $-1 < r_c \leq 1$; a value of 1 indicates complete positive correlation, and a value of 0 indicates the data sets are uncorrelated. The Ewald fluctuation results show significantly greater correlation with the experimentally derived fluctuations ($r_c = 0.68$ for 396 backbone heavy atoms; 0.63 for 362 side chain heavy atoms) than the 9/18 Å fluctuation results ($r_c = 0.49$, backbone; 0.49, side chain).

Energy conservation. Table II compares energy statistics for MD runs using the cutoff and PME methods. The cutoff methods show poor energy conservation (reflected by the average drift in total energy per step) caused by unstable numerical integration of the truncated Coulombic potential. This results in artificial heating of the system in the absence of temperature monitoring. Conversely, the PME method shows remarkably good energy conservation, with essentially no heating. This presumably results from the continuity of the Ewald pair potential calculated with the PME method.

IV. CONCLUSION

The Particle Mesh Ewald method⁸ is applied to the HIV-1 protease crystal⁹ using molecular dynamics. Striking improvement is obtained over truncated list methods. Simulation using a 9 Å cutoff did not converge in 300 ps. Inclusion of a 9/18 Å “twin-range” cutoff showed significant improvement. Simulation using the Ewald summation convention gave the best agreement with crystallographic

data. The results indicate that long-range electrostatic interactions are essential for proper representation of the HIV-1 PR crystal structure. The Ewald simulation reported here represents the first application of an efficient implementation of the Particle Mesh Ewald method,⁸ and illustrates the importance of including long-range electrostatic forces in large macromolecular systems.

ACKNOWLEDGMENTS

The authors thank Dr. A. Wlodawer for providing several crystallographic structures, and acknowledge grants of Supercomputer time on the Pittsburgh C90 from Cray Research.

^{a)} Author to whom correspondence should be addressed.

¹W. van Gunsteren and H. Berendsen, *Angew. Chem. Int. Ed. Engl.* **29**, 992 (1990).

²M. Belhadj, H. Alper, and R. Levy, *Chem. Phys. Lett.* **179**, 13 (1991).

³H. Schreiber and O. Steinhauser, *Biochemistry* **31**, 5856 (1992).

⁴H. Berendsen, in *Molecular Dynamics and Protein Structure*, edited by J. Hermans (Polycrystal Book Service, Western Springs, IL, 1985).

⁵P. Ewald, *Ann. Phys. (Leipzig)* **64**, 253 (1921).

⁶S. de Leeuw, J. Perram, and E. Smith, *Proc. R. Soc. London Ser. A* **373**, 27 (1980).

⁷E. Smith, *Proc. R. Soc. London Ser. A* **375**, 475 (1981).

⁸T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).

⁹A. Wlodawer, M. Miller, M. Jaskólski, B. Sathyanarayana, E. Baldwin, I. Weber, L. Selk, L. Clawson, J. Schneider, and S. Kent, *Science* **245**, 616 (1989).

¹⁰S. Weiner, P. Kollman, D. Case, U. Singh, C. Chio, G. Alagona, S. Profeta, and P. Weiner, *J. Am. Chem. Soc.* **106**, 765 (1984).

¹¹C. Foley, L. Pedersen, P. Charifson, T. Darden, A. Wittinghofer, E. Pai, and M. Anderson, *Biochemistry* **31**, 4951 (1992).

¹²S. Weiner and P. Kollman, *J. Comput. Chem.* **7**, 230 (1986).

¹³W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, and M. Klein, *J. Chem. Phys.* **79**, 926 (1983).

¹⁴D. York, T. Darden, L. Pedersen, and M. Anderson, *Biochemistry* **32**, 1443 (1993).

¹⁵J. Ryckaert, G. Ciccotti, and H. Berendsen, *J. Comput. Phys.* **23**, 327 (1977).

¹⁶M. Navia, P. Fitzgerald, B. McKeever, C-T Leu, J. Heimbach, W. Herber, I. Sigal, P. Darke, and J. Springer, *Nature* **37**, 615 (1989).

¹⁷G. Petsko and D. Ringe, *Ann. Rev. Biophys. Bioeng.* **13**, 331 (1984).

¹⁸W. Kabsch and C. Sander, *Biopolymers* **22**, 2577 (1983).