JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Accelerated Molecular Dynamics Simulations with the AMOEBA Polarizable Force Field on Graphics Processing Units

Steffen Lindert,*[,†,‡] Denis Bucher,[§,⊥] Peter Eastman,[∥] Vijay Pande,[∥] and J. Andrew McCammon[†,‡,§,⊥]

[†]Department of Pharmacology, University of California San Diego, La Jolla, California 92093United States

[‡]Center for Theoretical Biological Physics, La Jolla, California 92093United States

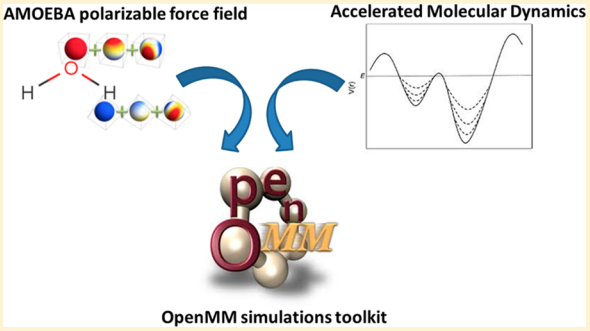
[§]Howard Hughes Medical Institute, University of California San Diego, La Jolla, California 92093United States

[∥]Department of Bioengineering, Stanford University, Stanford, California 94305, United States

[⊥]Department of Chemistry & Biochemistry, NSF Center for Theoretical Biological Physics, National Biomedical Computation Resource, University of California San Diego, La Jolla, California 92093, United States

Supporting Information

ABSTRACT: The accelerated molecular dynamics (aMD) method has recently been shown to enhance the sampling of biomolecules in molecular dynamics (MD) simulations, often by several orders of magnitude. Here, we describe an implementation of the aMD method for the OpenMM application layer that takes full advantage of graphics processing units (GPUs) computing. The aMD method is shown to work in combination with the AMOEBA polarizable force field (AMOEBA-aMD), allowing the simulation of long time-scale events with a polarizable force field. Benchmarks are provided to show that the AMOEBA-aMD method is efficiently implemented and produces accurate results in its standard parametrization. For the BPTI protein, we demonstrate that the protein structure described with AMOEBA remains stable even on the extended time scales accessed at high levels of accelerations. For the DNA repair metalloenzyme endonuclease IV, we show that the use of the AMOEBA force field is a significant improvement over fixed charged models for describing the enzyme active-site. The new AMOEBA-aMD method is publicly available (http://wiki.simtk.org/openmm/VirtualRepository) and promises to be interesting for studying complex systems that can benefit from both the use of a polarizable force field and enhanced sampling.



INTRODUCTION

Molecular dynamics (MD) is one of the most prominent techniques used to study the dynamics and equilibrium properties of biomolecules. It solves Newton's equations of motion for all atoms, using force fields that account for bonded and nonbonded atomic interactions. These force fields have been parametrized to agree with quantum mechanical simulations and experiments. Ever since the seminal first protein MD simulation of bovine pancreatic trypsin inhibitor (BPTI),[1] tremendous progress has been made both in terms of accuracy and sampling efficiency. Even larger macromolecular systems can now be simulated on the order of microseconds. However, many important biological processes occur on time scales far beyond this regime posing the need for millisecond or longer simulations. The development of the Anton supercomputer[2] certainly marked a large step in the right direction; however, that kind of computational fire power is not readily available to the average researcher. One of the most significant speed improvements of the last years was the utilization of graphics processing units (GPUs) for molecular dynamics. Efficient MD codes have been developed to aid with this endeavor.[3] Nonetheless, it will take many years before nonspecialized computer hardware can routinely simulate large molecular systems on the time scales of their slowest motions, which is needed to fully characterize their free energy landscapes.

In the meantime, enhanced sampling methods are one solution to this problem. A plethora of methods exist that perturb the underlying potential energy landscape to increase the chance of transitions between low energy states.[4] Three notable examples are the adaptive biasing force (ABF[5]), metadynamics[6] and driven adiabatic dynamics[7] methods. In addition to these methods, accelerated MD (aMD) is a promising technique that directly modifies regions of the potential energy landscape that are below a certain cutoff energy.[8] By energetically raising these regions and thus lowering barriers between energy wells, the landscape is perturbed to allow more frequent transitions between low energy states. This results in an enhanced sampling of the conformational space accessible to the simulation. aMD has been previously implemented for classical (nonpolarizable)

Received: June 17, 2013

   dx.doi.org/10.1021/ct400514p | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

simulations in AMBER[8a] and NAMD,[9] and for ab initio MD in CPMD.[10]

In addition to improvements in sampling (the precision problem), the force fields used to propagate atoms may also need to be improved (the accuracy problem). Moreover, sampling and force field accuracy are related challenges, since better sampling allows for more precise calculations of experimental quantities that in turn can lead to the detection of smaller force field deficiencies. Although the functional form of biomolecular force fields has remained largely unchanged since the first MD simulation of a protein conducted more than 35 years ago, recent studies highlight certain shortcomings present in all current force fields that are unlikely to be solved simply by tuning their parametrization. For example, none of the current methods seem to be able to accurately capture the temperature dependency of proteins secondary structure propensities.[11] In addition, polarizability may be essential to accurately simulate highly charged systems, such as nucleic acids in DNA and RNA.[12] For this reason, substantial efforts are under way to develop force fields with more sophisticated functional forms, including polarizable force fields that can capture the redistribution of electrons around each atom in response to changes in the environment. This approach may eventually replace conventional force fields in several key areas (e.g., the modeling of protein folding,[13] computer-aided drug design,[14] the calculation of ion channel properties,[15] or the description of allostery[16]).

The AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications) polarizable force field[17] was developed by Ponder and co-workers with the aim to move away from the well-established fixed point charge models and toward more expensive models that should allow a more accurate description of molecular properties. AMOEBA is one of the most widely used polarizable force fields to date, and it has been demonstrated to perform better than nonpolarizable force fields for describing solvation free energies of drug-like small molecules, and dynamical properties away from ambient condition.[18] However, despite the impressive acceleration provided by GPU computing,[19] the time performance of AMOEBA remains a significant limitation to its wider use to simulate large-scale biomolecular systems, since such simulations remain about 1−2 orders of magnitude more expensive than their nonpolarizable equivalents. For this reason, practical usage of AMOEBA has been limited, especially when target events are in micro- to millisecond time scales. AMOEBA has been previously combined with an enhanced-sampling free energy method—orthogonal space random walk (OSRW)—to achieve ab initio prediction of organic crystal structures and thermodynamics.[20] Similarly, combining AMOEBA with the enhanced sampling capabilities of aMD represents an attractive prospect. OpenMM is the perfect platform to implement the aMD method, as it already allows a wide user base to take advantage of GPU-accelerated MD simulations.[3c,19]

In this paper, we show that the aMD method and the AMOEBA force field can be synergistically combined in a way that conserves the accuracy of the polarizable force field, while significantly enhancing sampling. Using three examples, we show that (1) with larger energy boost levels, the AMEOBA-based aMD simulations can maintain structural stability, (2) the choice of aMD parameters is the same for AMOEBA and for nonpolarizable force fields, and (3) the AMOEBA-based aMD simulations allow metalloenzyme reactive sites to be better simulated.

MATERIAL AND METHODS

Implementation of aMD into OpenMM. Over the years, different variants of the aMD algorithm have been proposed.[8b] In the original implementation,[8a] aMD was used to only boost the dihedral potential since many protein conformational changes are governed by changes in the torsional degrees of freedom. Subsequent work introduced an aMD version that boosted the total potential.[21] In a more recent implementation, aMD uses a dual boosting approach that applies one boost potential only to the torsional terms and another separate boost potential to the total potential.[22] Here, we present the implementation of all three flavors of aMD into OpenMM using the CustomIntegrator, an integrator used to implement arbitrary, user defined integration algorithms. A detailed summary of the theory of aMD can be found in refs 8a and 22, so that we will only list the most important equations that were used for the implementation. In aMD, whenever the system's potential energy, $V(\mathbf{r})$, falls below a threshold energy, E, a boost potential, $\Delta V(\mathbf{r})$, is added to yield the biased potential $V^*(\mathbf{r})$.

$$V^*(\mathbf{r}) = \mathrm{V}(\mathbf{r}) + \Delta V(\mathbf{r}) = V(\mathbf{r}) + \begin{cases} 0, & V(\mathbf{r}) \geq E \\ \dfrac{(E - \mathrm{V}(\mathbf{r}))^2}{\alpha + E - V(\mathbf{r})}, & V(\mathbf{r}) < E \end{cases}$$

Introducing the following abbreviations

$$\beta \equiv E - V(\mathbf{r})$$

$$\gamma \equiv \frac{\alpha^2}{(\alpha + \beta)^2}$$

The total force on the system is calculated as a sum of constitutive forces, each obtained from the different components ("comp") of the potential energy (e.g., dihedrals). A general expression can be derived for the aMD forces that is valid irrespectively of the choice made for the boosted components, as[8a,22]

$$F_{\mathrm{comp}}{}^* = -\nabla V_{\mathrm{comp}}(\mathbf{r}) \frac{\alpha_{\mathrm{comp}}{}^2}{(\alpha_{\mathrm{comp}} + \beta_{\mathrm{comp}})^2} = F_{\mathrm{comp}}\gamma_{\mathrm{comp}}$$

The overall force in the boosted system F^* is then obtained as

$$F^* = (F - F_{\mathrm{comp}}) + F_{\mathrm{comp}}\gamma_{\mathrm{comp}}$$

where F is the total force of the unaccelerated system and F_{comp} is the unaccelerated force of the component that is accelerated (e.g., dihedrals). The forces for the three flavors of aMD then become

dihedral boost:

$$F^* = (F - F_{\mathrm{torsion}}) + F_{\mathrm{torsion}}\gamma_{\mathrm{torsion}} \quad \text{with} \quad \gamma_{\mathrm{torsion}} \equiv \frac{\alpha_{\mathrm{torsion}}{}^2}{(\alpha_{\mathrm{torsion}} + \beta_{\mathrm{torsion}})^2}$$

total energy boost:

$$F^* = (F - F_{\mathrm{tot}}) + F_{\mathrm{tot}}\gamma_{\mathrm{tot}} = F\gamma_{\mathrm{tot}} \quad \text{since} \quad F_{\mathrm{tot}} \equiv F$$

dx.doi.org/10.1021/ct400514p | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. Acceleration Parameters for the aMD Simulations

system	force field	α_{tot} [kcal/mol]	E_{tot} [kcal/mol]	$\alpha_{torsion}$ [kcal/mol]	$E_{torsion}$ [kcal/mol]
alanine dipeptide	ff99SBildn	382.4	−5668.2	2.1	19.5
BPTI	ff99SBildn	3431.0	−50 335.4	40.6	821.3
	AMOEBA	3431.0	−56 857.9	40.6	310.1
endonuclease IV	ff99SBildn	10 402.2	−149 530.1	199.5	4,129.6
	AMOEBA	10 402.2	−169 944.8	199.5	1,971.8

dual boost:

$$F^* = (F - F_{torsion})\gamma_{tot} + F_{torsion}\gamma_{torsion}$$

Note that, in the dual boost implementation, the dihedral force is only boosted once. This is in agreement with current aMD implementations in AMBER and NAMD. The Custom Integrator allows convenient implementation of aMD, by scaling the forces at every simulation step according to the boost. For this, the energy and forces associated with the component to be boosted have to be retrieved. Then, the total force is scaled depending on whether the system is above or below the threshold energy. The aMD CustomIntegrator is available online at http://wiki.simtk.org/openmm/VirtualRepository.

Systems and MD Simulations. Three different systems were studied: the alanine dipeptide (N-acetyl-N′-methyl-alaninamide), BPTI, and endonuclease IV. Alanine dipeptide was solvated with a water box containing a buffer region of 10 Å. To eliminate any steric clashes, 100 steps of conjugate gradient minimization were performed using SANDER.[23] A short 50 ps NPT simulation was used to heat up the system to a temperature of 300 K. BPTI and Endonuclease IV were prepared based on PDB entries 5PTI[24] and 1QTW,[25] respectively. Tleap[23] was used to neutralize the systems by adding 6 Cl^- and 6 Na^+ counterions respectively and solvate them with a water box. TIP3P water[26] was used for the nonpolarizable simulations. The fully solvated system contained 17 155 and 52 011 atoms, respectively. Minimization using SANDER[23] was carried out in two stages: an initial minimization of 1000 steps for the solvent and ions with the protein restrained by a force constant of 500 kcal/mol/Å^2, followed by a 2500 step minimization of the entire system. A short initial 20 ps MD simulation with weak restraints (10 kcal/mol/Å^2) on the protein residues was used to heat the system to a temperature of 300 K.

All production MD simulations for the alanine dipeptide, BPTI, and endonuclease IV were conducted in the NPT ensemble at 300 K, using the OpenMM Python-based application layer.[19] Periodic boundary conditions were used, along with a nonbonded interaction cutoff of 10 Å. Particle Mesh Ewald was used.[27] Nonpolarizable simulations were performed using the AMBER ff99SBildn force field[28] while polarizable simulations were performed using the AMOEBA force field.[18] For the AMBER simulations, bonds involving hydrogen atoms were constrained using the SHAKE algorithm,[29] allowing for a time step of 2 fs. A time step of 1 fs was used for the AMOEBA simulations, while no constraints were used. Mutual polarization was used along with an induced target ϵ of 0.01. Energy conservation was monitored and none of the simulations showed energy drift.

The acceleration level for the aMD method is defined in terms of E and α, where E is the threshold boost energy and α is a tuning parameter that determines the shape of the accelerated potential.[8a] To define the acceleration parameters, equilibrium MD simulations were first conducted, with the AMBER ff99SBildn and AMOEBA force fields, to calculate the average potential and dihedral energies. For the total potential/torsional energy term, the boost energy was defined as the average total/dihedral angle energy plus 3.5 times the number of residues in the solute, and α was defined as 20% of E. From previous experience, this empirical formula produces marked enhancement in the sampling, while allowing for the accurate reconstruction of the thermodynamic ensemble.[9,30] Despite the explicit inclusion of polarization and anharmonic corrections in the AMOEBA force field, we have found that similar parameters can be used to enhance the sampling in AMBER ff99SBildn and AMOEBA aMD simulations. Table 1 summarizes the boost parameters for our aMD simulations.

RESULTS AND DISCUSSION

Our results are presented as follows; first we use the workhorse of molecular simulations, the alanine dipeptide, to ensure that the aMD implementation does indeed significantly enhance the conformational space sampling. We then use BPTI to show that even at very high acceleration levels, AMOEBA-aMD simulations are stable, as indicated by low root-mean-square deviations (RMSD) with respect to the BPTI crystal structure. Finally, we study the enzyme endonuclease IV to showcase a large protein system, for which nonpolarizable simulations fail to accurately describe structural details.

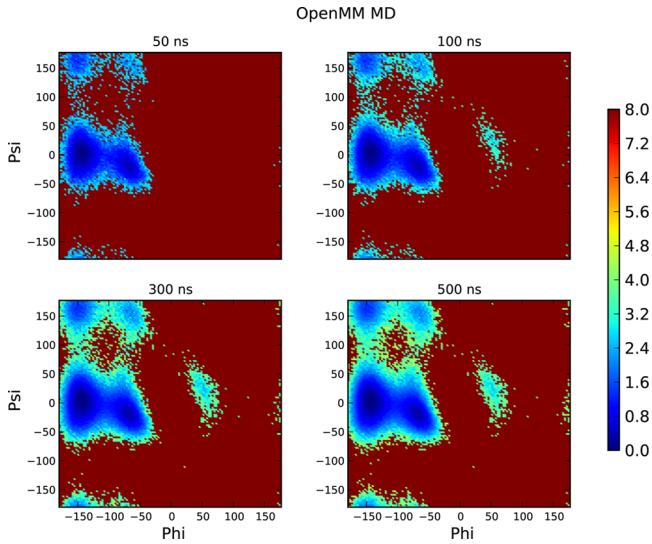


Figure 1. $\varphi-\psi$ free energy maps (in kcal/mol) for alanine dipeptide calculated from cMD simulations of length (A) 50 ns, (B) 100 ns, (C) 300 ns, and (D) 500 ns.

aMD Samples All Conformations of the Alanine Dipeptide. To test our aMD implementation, four different types of simulations were performed for alanine dipeptide: 500 ns of conventional MD (cMD), 100 ns of dihedral boost aMD,

dx.doi.org/10.1021/ct400514p | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

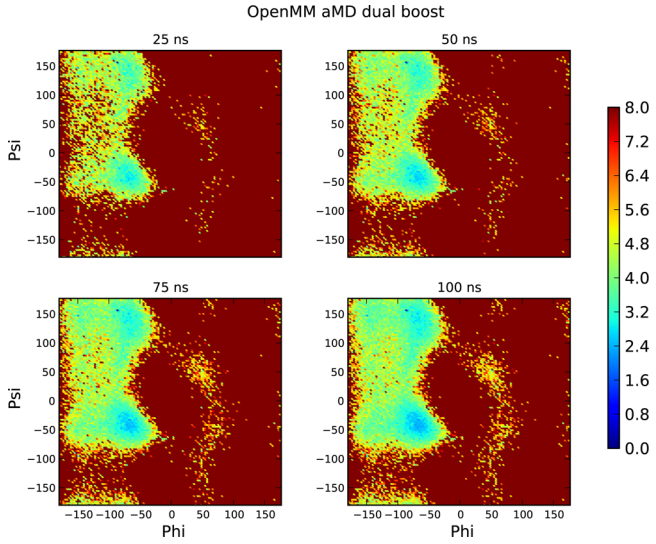


Figure 2. Reweighted $\varphi-\psi$ free energy maps (in kcal/mol) for alanine dipeptide calculated from dual boost aMD simulations of length (A) 25 ns, (B) 50 ns, (C) 75 ns, and (D) 100 ns.

100 ns of total energy boost aMD, and 100 ns of dual boost aMD. The conformational space sampling is evaluated by deriving a free energy profile from occupancy of states in the $\varphi-\psi$ dihedral backbone angle space. Pioneering theoretical work in the late 1980s and early 1990s established the four free energy minima accessible to alanine dipeptide in the $\varphi-\psi$ conformational space:[31] the β region ($\varphi < 0°$ and $120° < \psi < 180°$, corresponding to a β-strand conformation), the α_R region ($\varphi < 0°$ and $-60° < \psi < 60°$, corresponding to a right-handed α-helical conformation), the α_L region ($\varphi \sim 60°$ and $\psi \sim 60°$, corresponding to a left-handed α-helical conformation) and the C_7^{ax} region ($\varphi \sim 60°$ and $\psi \sim -80°$). For an excellent graphical representation of the free energy surface please refer to ref 32.

Figure 1 shows the $\varphi-\psi$ free energy maps calculated from the cMD simulation and different simulation times. Despite the simplicity of the alanine dipeptide system, conventional MD cannot sample the entire $\varphi-\psi$ dihedral backbone angle space. Within the first 50 ns, only the two deepest energy minima (β and α_R) are sampled. Within 500 ns sampling is extended to cover roughly the α_L region. No sampling of the C_7^{ax} region is seen within 500 ns of cMD simulation.

In contrast, dual boost aMD samples the conformational space much more efficiently. Figure 2 shows the $\varphi-\psi$ free energy maps for different simulation times of the dual boost aMD simulation. Even within 25 ns, all four regions (including the C_7^{ax} region) are sampled. After 100 ns of dual boost aMD the reweighted energy profile clearly matches previously reported profile generated by umbrella sampling.[32] This demonstrates the power of the aMD method to enhance sampling by orders of magnitude and its correct implementation in OpenMM. It is worth noting that an explicit boost of the dihedral energy is often beneficial for peptide and protein

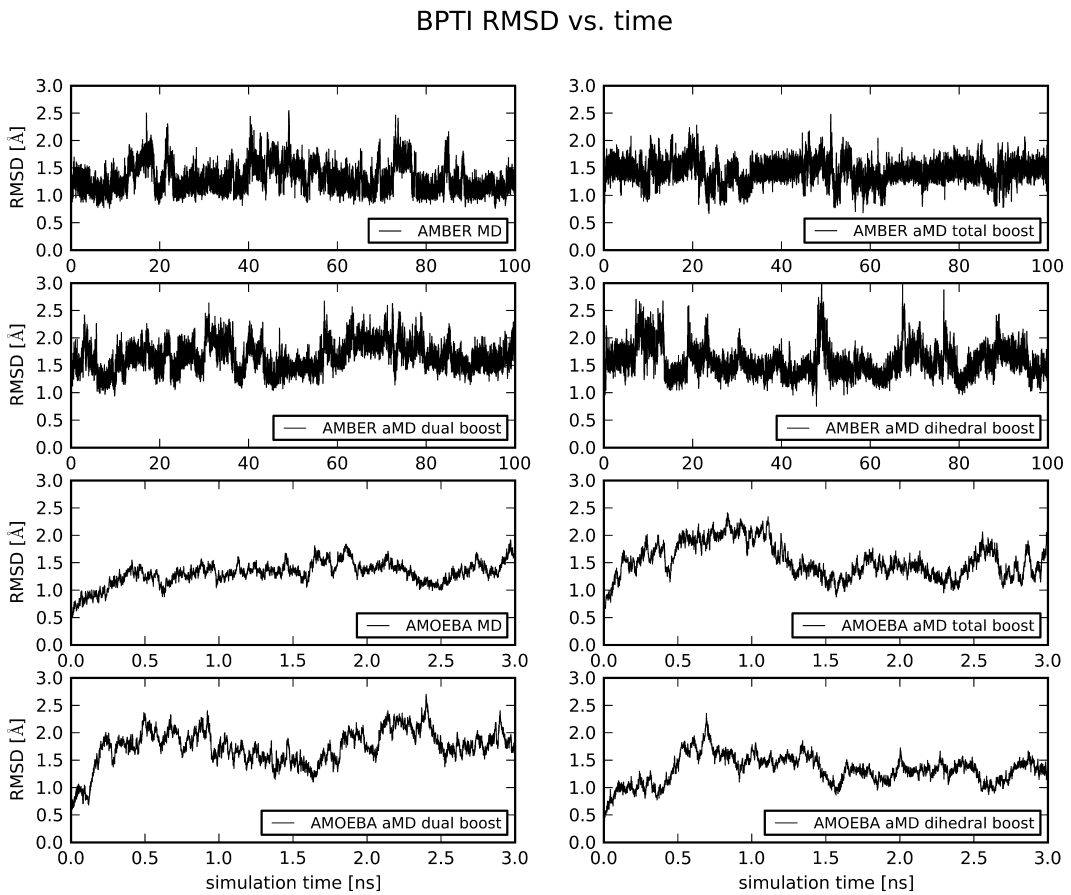


Figure 3. RMSD vs time plots for simulations of BPTI. RMSD values over the course of the simulations are shown for eight different BPTI simulations: MD AMBER ff99SBildn, total energy boost aMD AMBER ff99SBildn, dual boost aMD AMBER ff99SBildn, dihedral boost aMD AMBER ff99SBildn, MD AMOEBA, total energy boost aMD AMOEBA, dual boost aMD AMOEBA, and dihedral boost aMD AMOEBA.

dx.doi.org/10.1021/ct400514p | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

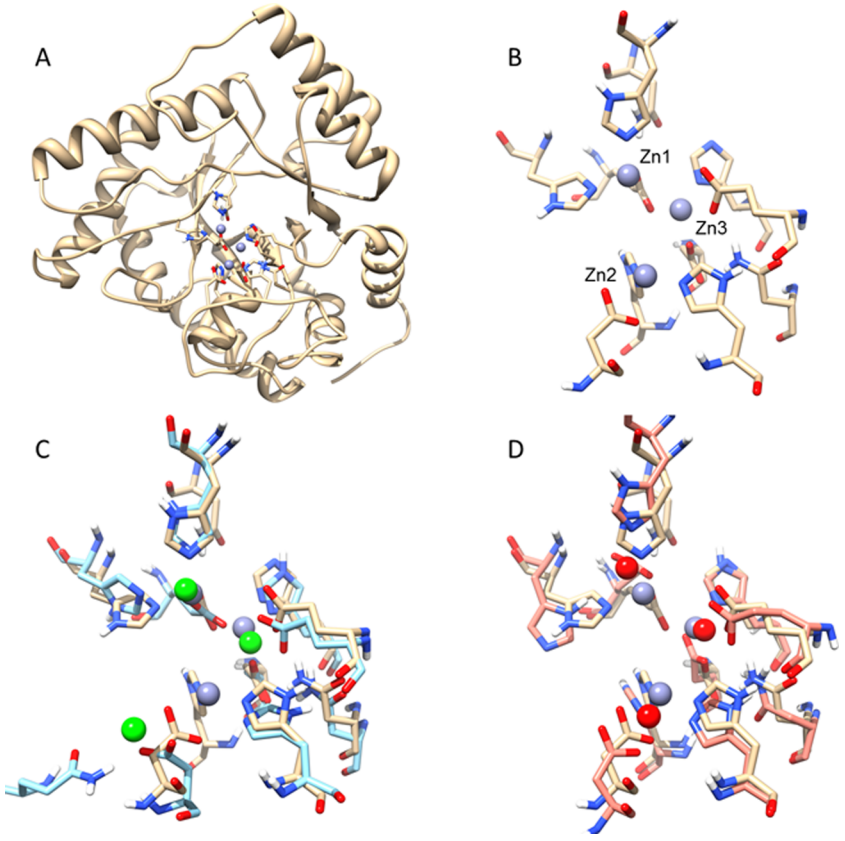


Figure 4. Structure of endonuclease IV. (A) Ribbon representation of entire protein with its zinc cluster in the center. (B) Close-up view of the endonuclease IV active site in the high resolution crystal structure (1QTW, 1.02 Å resolution). The zinc ions are forming a right triangle geometry coordinated by histidine, aspartate, and glutamate residues. (C) Superposition of the crystal structure active site and the active site after 3 ns of AMBER ff99SBildn simulation. Crystal structure zinc is shown in gray, while the simulated zinc is shown in green. (D) Superposition of the crystal structure active site and the active site after 3 ns of AMOEBA simulation. Crystal structure zinc is shown in gray, while the simulated zinc is shown in red.

simulations. For instance, in this test, only boosting the total energy was not sufficient to sample the α_L and C_7^{ax} regions within 100 ns (Supporting Information Figure 1), while the dihedral boost aMD samples the conformational space similar to the dual boost aMD (Supporting Information Figure 2).

AMOEBA-aMD Simulations of BPTI Do Not Exhibit Instabilities. The BPTI protein is often used to test new methods. In particular, it was used to test a previous implementation of aMD for the Amber code,[33] based on nonpolarizable AMBER force fields. It was shown that, at high acceleration levels, aMD simulations of ∼500 ns can replicate all the energy minima of BPTI that were observed by Shaw and co-workers using millisecond simulations,[34] thus leading to an acceleration in the sampling of 3−4 orders of magnitude.

Here, we were interested to show that BPTI is stable at similar high acceleration levels in our aMD OpenMM implementation, with both the AMBER ff99SBildn, and the AMOEBA force fields. For each force field, a total of four different simulations were performed: cMD, dihedral boost aMD, total energy boost aMD and dual boost aMD. The simulations were run for 100 ns, and 3 ns, for the AMBER ff99SBildn and AMOEBA force fields, respectively. Simulation speeds of ∼30 ns/day for AMBER and ∼300 ps/day for AMOEBA were observed. To allow for a fair comparison, we also performed a short 1 ns benchmark AMOEBA simulation using the RESPA integrator (allowing for a 2 fs time step for the multipole force evaluation, while maintaining a 1 fs time step for all other forces). The simulations ran at 632 ps/day. Thus, depending on the parameters, AMOEBA simulations are about 1−2 orders of magnitude slower than cMD; however, enhanced sampling with aMD has been previously shown for BPTI to enhance sampling by several orders of magnitude[33] and therefore should be able to bridge this gap. The BPTI protein stability was judged by its RMSD to the starting structure. Potential unfolding events or instabilities due to excessive boosting would show up as an increase in the RMSD. Figure 3 summarizes the RMSDs for the simulations and shows that none of the aMD simulations exceeds 3 Å in the RMSD, which is similar to the unaccelerated simulations.

AMOEBA Improves the Description of Endonuclease IV. Next, we demonstrate that AMOEBA can achieve more accurate results than nonpolarizable force field simulations. For this, we chose to simulate the endonuclease IV metallo-enzyme. The enzyme has been determined at a very high resolution (1.02 Å) using X-ray crystallography.[25] Figure 4A,B shows the enzyme with its active site. It contains a stable three zinc cluster coordinated by histidine, aspartate, and glutamate residues. The zinc ions form a structure reminiscent of a right triangle with distinct relative distances. The trinuclear cluster has been previously studied with hybrid quantum mechanics/molecular mechanics (QM/MM) simulations and was found to be stable.[35] However, QM/MM simulations are typically time-consuming and can often probe only the picosecond time scale. Our initial hypothesis was that AMOEBA simulations should be

dx.doi.org/10.1021/ct400514p | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

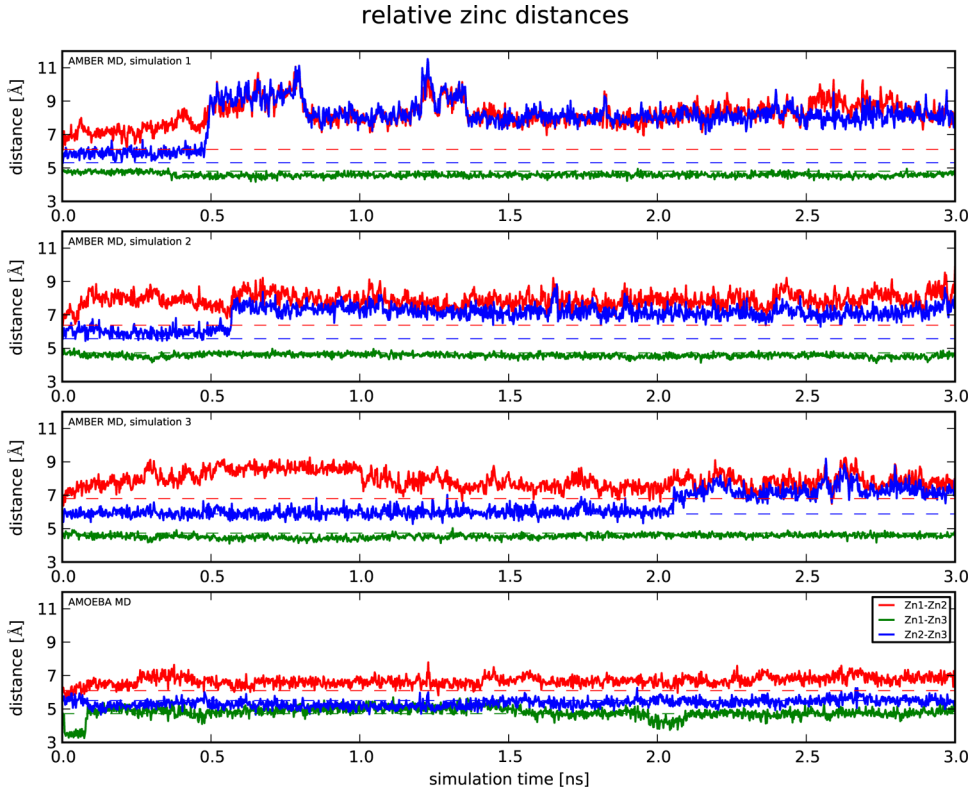


Figure 5. Zn−Zn distances in endonuclease IV over the course of the simulations. Endonuclease IV contains a three zinc cluster (Zn 286, 287, and 288). Distances between 286 and 287 (red), 286−288 (green), and 287−288 (blue) are shown. Horizontal dotted colored lines indicate the distance between the respective ion pair in the high-resolution crystal structure. Three independent MD simulations with the AMBER ff99SBildn force field are shown on top. A simulation with the AMOEBA force field is shown in the bottom.

able to more accurately describe the zinc cluster geometry than nonpolarizable AMBER simulations. To show this, we ran nonpolarizable and polarizable simulations on the protein and monitored the relative zinc distances. Again, depending on the simulation parameters, we observed about 1−2 orders of magnitude faster sampling for nonpolarizable force field simulations (~9 ns/day, for AMBER, versus ~90 ps/day, for AMOEBA without RESPA). However, the accuracy was improved with AMOEBA. For instance, Figure 5 shows the Zn−Zn distances over the course of the simulations. In all three independent AMBER ff99SBildn simulations, the relative positions of the zinc cluster were perturbed within 3 ns, demonstrating a failure of the nonpolarizable force field to correctly capture the coordination structure around these charged ions. In contrast, the AMOEBA force field maintained the zinc geometry over the course of the 3 ns simulation. Figure 4C,D summarizes the active site residues and zinc positions after 3 ns of simulation, superimposed with the crystal structure. The most pronounced perturbation in the nonpolarizable simulations is the displacement of Zn2 away from the active site, facilitated by the displacement of the coordinating side chain ASP 229, and transferring the coordination of Zn2 from HIS 182 to GLN 150.

CONCLUSIONS

In this paper, we presented the synergistic combination of the enhanced sampling method aMD with the polarizable force field AMOEBA. The implementation of the aMD sampling method into OpenMM should provide a fast and simple avenue for studying biomolecular systems that require extensive sampling (e.g., micro- to milliseconds), in addition to a polarizable force field. The tests presented here suggest that it is possible to maintain the accuracy of the AMOEBA polarizable force field while significantly improving the simulation sampling efficiency. At the acceleration levels used, the sampling efficiency for medium size protein simulations is typically increased by 2−3 orders in magnitude.[33,36] Finally, the superior performance of AMOEBA was demonstrated for describing the active-site of a zinc metalloenzyme, which is consistent with previous studies.[37] The AMOEBA-aMD implementation takes full advantage of GPU computing and is publicly available in OpenMM (https://simtk.org/home/openmm).

ASSOCIATED CONTENT

Supporting Information

Reweighted $\varphi-\psi$ free energy maps (Figures S1 and S2). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: 858-534-2913. Fax: 858-534-4974. E-mail: slindert@ucsd.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was made possible by a Simbios Visiting Scholar Fellowship awarded to S.L. (U54 GM072970). Furthermore, it was supported by the National Institutes of Health, the

National Science Foundation, the Howard Hughes Medical Institute, the National Biomedical Computation Resource, and the National Science Foundation (NSF) Supercomputer Centers. Computational resources were supported, in part, by the National Science Foundation grant PHY-0822283, the Center for Theoretical Biological Physics. S.L. was supported by the American Heart Association and the Center for Theoretical Biological Physics.

REFERENCES

(1) McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of folded proteins. *Nature* **1977**, *267* (5612), 585−590.
(2) Shaw, D. E.; Deneroff, M. M.; Dror, R. O.; Kuskin, J. S.; Larson, R. H.; Salmon, J. K.; Young, C.; Batson, B.; Bowers, K. J.; Chao, J. C.; Eastwood, M. P.; Gagliardo, J.; Grossman, J. P.; Ho, C. R.; Ierardi, D. J.; Kolossvary, I.; Klepeis, J. L.; Layman, T.; Mcleavey, C.; Moraes, M. A.; Mueller, R.; Priest, E. C.; Shan, Y. B.; Spengler, J.; Theobald, M.; Towles, B.; Wang, S. C. Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM* **2008**, *51* (7), 91−97.
(3) (a) Gotz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. *J. Chem. Theory Comput.* **2012**, *8* (5), 1542−1555. (b) Stone, J. E.; Hardy, D. J.; Ufimtsev, I. S.; Schulten, K. GPU-accelerated molecular modeling coming of age. *J. Mol. Graph Modell.* **2010**, *29* (2), 116−125. (c) Eastman, P.; Pande, V. S. OpenMM: A hardware-independent framework for molecular simulations. *Comput Sci Eng.* **2010**, *12* (4), 34−39.
(4) (a) Sinko, W.; Lindert, S.; McCammon, J. A. Accounting for receptor flexibility and enhanced sampling methods in computer-aided drug design. *Chem. Biol. Drug Des.* **2013**, *81* (1), 41−49. (b) Lei, H.; Duan, Y. Improved sampling methods for molecular simulation. *Curr. Opin. Struct. Biol.* **2007**, *17* (2), 187−191. (c) Mitsutake, A.; Mori, Y.; Okamoto, Y. Enhanced sampling algorithms. *Methods Mol. Biol.* **2013**, *924*, 153−195.
(5) Darve, E.; Pohorille, A. Calculating free energies using average force. *J. Chem. Phys.* **2001**, *115* (20), 9169−9183.
(6) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (20), 12562−12566.
(7) Cuendet, M. A.; Tuckerman, M. E. Alchemical free energy differences in flexible molecules from thermodynamic integration or free energy perturbation combined with driven adiabatic dynamics. *J. Chem. Theory Comput.* **2012**, *8* (10), 3504−3512.
(8) (a) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **2004**, *120* (24), 11919−11929. (b) Markwick, P. R.; McCammon, J. A. Studying functional dynamics in bio-molecules using accelerated molecular dynamics. *Phys. Chem. Chem. Phys.* **2011**, *13* (45), 20053−20065.
(9) Wang, Y.; Harrison, C. B.; Schulten, K.; McCammon, J. A. Implementation of accelerated molecular dynamics in NAMD. *Comput. Sci. Discov.* **2011**, *4* (1), 015002.
(10) Bucher, D.; Pierce, L. C.; McCammon, J. A.; Markwick, P. R. On the use of accelerated molecular dynamics to enhance configurational sampling in ab initio simulations. *J. Chem. Theory Comput.* **2011**, *7* (4), 890−897.
(11) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Systematic validation of protein force fields against experimental data. *PLoS One* **2012**, *7* (2), e32131.
(12) Baker, C. M.; Anisimov, V. M.; MacKerell, A. D., Jr. Development of CHARMM polarizable force field for nucleic acid bases based on the classical Drude oscillator model. *J. Phys. Chem. B* **2011**, *115* (3), 580−596.
(13) van der Vaart, A.; Bursulaya, B. D.; Brooks, C. L.; Merz, K. M. Are many-body effects important in protein folding? *J. Phys. Chem. B* **2000**, *104* (40), 9554−9563.
(14) Cho, A. E.; Guallar, V.; Berne, B. J.; Friesner, R. Importance of accurate charges in molecular docking: Quantum mechanical/molecular mechanical (QM/MM) approach. *J. Comput. Chem.* **2005**, *26* (9), 915−931.
(15) Bucher, D.; Raugei, S.; Guidoni, L.; Dal Peraro, M.; Rothlisberger, U.; Carloni, P.; Klein, M. L. Polarization effects and charge transfer in the KcsA potassium channel. *Biophys. Chem.* **2006**, *124* (3), 292−301.
(16) Ionescu, C. M.; Svobodova Varekova, R.; Prehn, J. H.; Huber, H. J.; Koca, J. Charge profile analysis reveals that activation of pro-apoptotic regulators Bax and Bak relies on charge transfer mediated allosteric regulation. *PLoS Comput. Biol.* **2012**, *8* (6), e1002565.
(17) Ren, P.; Ponder, J. W. Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J. Comput. Chem.* **2002**, *23* (16), 1497−1506.
(18) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.; Head-Gordon, M.; Clark, G. N.; Johnson, M. E.; Head-Gordon, T. Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B* **2010**, *114* (8), 2549−2564.
(19) Eastman, P.; Friedrichs, M. S.; Chodera, J. D.; Radmer, R. J.; Bruns, C. M.; Ku, J. P.; Beauchamp, K. A.; Lane, T. J.; Wang, L. P.; Shukla, D.; Tye, T.; Houston, M.; Stich, T.; Klein, C.; Shirts, M. R.; Pande, V. S. OpenMM 4: A reusable, extensible, hardware-independent library for high performance molecular simulation. *J. Chem. Theory Comput.* **2013**, *9* (1), 461−469.
(20) Schnieders, M. J.; Baltrusaitis, J.; Shi, Y.; Chattree, G.; Zheng, L.; Yang, W.; Ren, P. The structure, thermodynamics and solubility of organic crystals from simulation with a polarizable force field. *J. Chem. Theory Comput.* **2012**, *8* (5), 1721−1736.
(21) de Oliveira, C. A.; Hamelberg, D.; McCammon, J. A. On the application of accelerated molecular dynamics to liquid water simulations. *J. Phys. Chem. B* **2006**, *110* (45), 22695−22701.
(22) Hamelberg, D.; de Oliveira, C. A.; McCammon, J. A. Sampling of slow diffusive conformational transitions with accelerated molecular dynamics. *J. Chem. Phys.* **2007**, *127* (15), 155102.
(23) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The AMBER biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26* (16), 1668−1688.
(24) Wlodawer, A.; Walter, J.; Huber, R.; Sjolin, L. Structure of bovine pancreatic trypsin inhibitor. Results of joint neutron and X-ray refinement of crystal form II. *J. Mol. Biol.* **1984**, *180* (2), 301−329.
(25) Hosfield, D. J.; Guan, Y.; Haas, B. J.; Cunningham, R. P.; Tainer, J. A. Structure of the DNA repair enzyme endonuclease IV and its DNA complex: Double-nucleotide flipping at abasic sites and three-metal-ion catalysis. *Cell* **1999**, *98* (3), 397−408.
(26) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926−935.
(27) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald—An N.Log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.
(28) (a) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65* (3), 712−725. (b) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the AMBER ff99SB protein force field. *Proteins* **2010**, *78* (8), 1950−1958.
(29) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327−341.
(30) (a) Lindert, S.; Kekenes-Huskey, P. M.; Huber, G.; Pierce, L.; McCammon, J. A. Dynamics and calcium association to the N-terminal regulatory domain of human cardiac troponin C: A multiscale computational study. *J. Phys. Chem. B* **2012**, *116* (29), 8449−8459. (b) Bucher, D.; Grant, B. J.; Markwick, P. R.; McCammon, J. A. Accessing a hidden conformation of the maltose binding protein using

accelerated molecular dynamics. *PLoS Comput. Biol.* **2011**, *7* (4), e1002034.
(31) (a) Anderson, A. G.; Hermans, J. Microfolding: Conformational probability map for the alanine dipeptide in water from molecular dynamics simulations. *Proteins* **1988**, 3 (4), 262−265. (b) Tobias, D. J.; Brooks, C. L. Conformational equilibrium in the alanine dipeptide in the gas phase and aqueous solution—A comparison of theoretical results. *J. Phys. Chem.* **1992**, *96* (9), 3864−3870.
(32) Chekmarev, D. S.; Ishida, T.; Levy, R. M. Long-time conformational transitions of alanine dipeptide in aqueous solution: Continuous and discrete-state kinetic models. *J. Phys. Chem. B* **2004**, *108* (50), 19487−19495.
(33) Pierce, L. C.; Salomon-Ferrer, R.; Augusto, F. d. O. C.; McCammon, J. A.; Walker, R. C. Routine access to millisecond time scale events with accelerated molecular dynamics. *J. Chem. Theory Comput.* **2012**, *8* (9), 2997−3002.
(34) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. Atomic-level characterization of the structural dynamics of proteins. *Science* **2010**, *330* (6002), 341−346.
(35) (a) Ivanov, I.; Tainer, J. A.; McCammon, J. A. Unraveling the three-metal−ion catalytic mechanism of the DNA repair enzyme endonuclease IV. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (5), 1465−1470. (b) Bucher, D.; Masson, F.; Arey, J. S.; Röthlisberger, U. In *Quantum Biochemistry*; Wiley-VCH Verlag GmbH & Co. KGaA: Berlin, 2009; pp 517−535.
(36) Bucher, D.; Grant, B. J.; Markwick, P. R.; McCammon, J. A. Accessing a hidden conformation of the maltose binding protein using accelerated molecular dynamics. *PLoS Comput Biol.* **2011**, *7* (4), e1002034 DOI: doi:10.1371/journal.pcbi.1002034.
(37) (a) Wu, J. C.; Piquemal, J. P.; Chaudret, R.; Reinhardt, P.; Ren, P. Polarizable molecular dynamics simulation of Zn(II) in water using the AMOEBA force field. *J. Chem. Theory Comput.* **2010**, *6* (7), 2059−2070. (b) Xiang, J. Y.; Ponder, J. W. A valence bond model for aqueous Cu(II) and Zn(II) ions in the AMOEBA polarizable force field. *J. Comput. Chem.* **2013**, *34* (9), 739−749.