TIGER2 with solvent energy averaging (TIGER2A): An accelerated sampling method for large molecular systems with explicit representation of solvent

Xianfeng Li,¹ James A. Snyder,¹ Steven J. Stuart,² and Robert A. Latour^{1,a)}

¹Department of Bioengineering, Clemson University, Clemson, South Carolina 29634, USA

²Department of Chemistry, Clemson University, Clemson, South Carolina 29634, USA

(Received 25 June 2015; accepted 22 September 2015; published online 9 October 2015)

The recently developed “temperature intervals with global exchange of replicas” (TIGER2) accelerated sampling method is found to have inaccuracies when applied to systems with explicit solvation. This inaccuracy is due to the energy fluctuations of the solvent, which cause the sampling method to be less sensitive to the energy fluctuations of the solute. In the present work, the problem of the TIGER2 method is addressed in detail and a modification to the sampling method is introduced to correct this problem. The modified method is called “TIGER2 with solvent energy averaging,” or TIGER2A. This new method overcomes the sampling problem with the TIGER2 algorithm and is able to closely approximate Boltzmann-weighted sampling of molecular systems with explicit solvation. The difference in performance between the TIGER2 and TIGER2A methods is demonstrated by comparing them against analytical results for simple one-dimensional models, against replica exchange molecular dynamics (REMD) simulations for sampling the conformation of alanine dipeptide and the folding behavior of (AAQAA)₃ peptide in aqueous solution, and by comparing their performance in sampling the behavior of hen egg-white lysozyme in aqueous solution. The new TIGER2A method solves the problem caused by solvent energy fluctuations in TIGER2 while maintaining the two important characteristics of TIGER2, i.e., (1) using multiple replicas sampled at different temperature levels to help systems efficiently escape from local potential energy minima and (2) enabling the number of replicas used for a simulation to be independent of the size of the molecular system, thus providing an accelerated sampling method that can be used to efficiently sample systems considered too large for the application of conventional temperature REMD. © 2015 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4932341>]

I. INTRODUCTION

The “temperature intervals with global exchange of replicas” (TIGER2) method^{1,2} was recently developed to overcome some limitations of the conventional “replica exchange molecular dynamics” (REMD) method,³ which becomes increasingly computationally expensive as the size of the molecular system being simulated increases. For complex systems with many degrees of freedom, the REMD method requires the use of a large number of replicas at closely spaced temperature levels in order that the overlap between potential energy distributions of adjacent replicas is sufficient to obtain an acceptable exchange rate. As a consequence, in addition to the high computational cost, the large number of replicas results in slow diffusion of replicas over the numerous temperature levels, thus resulting in poor sampling efficiency. The TIGER2 method adopts the REMD approach of using multiple replicas sampled at different temperatures, using thermal fluctuations to allow the system to escape from local potential energy minima, but was specifically designed for systems that are too large for conventional REMD. It is able to conduct the simulation of large molecular systems with greater computational efficiency and much lower computational cost compared to REMD.

Similar to REMD, a TIGER2 simulation^{1,2} involves the use of a number of replicas (N_r) running in parallel at different

temperature levels. One of these replicas, the baseline replica, is sampled using ordinary molecular dynamics (MD) at the target temperature level, and the remaining $N_r - 1$ replicas undergo heating–sampling–quenching stages and so are called quenched replicas. During these stages, each quenched replica is heated to an elevated temperature and equilibrated, and then sampled at that temperature level using MD so that thermal fluctuations can help the systems cross over energy barriers. After sampling at the designated elevated temperature levels, all replicas are quenched down to the baseline temperature and thermally equilibrated. One of the quenched replicas is then randomly selected and compared to the baseline replica using the Boltzmann (Metropolis) importance sampling criterion,⁴ so that the new state is accepted with probability

$$P = \min(1, e^{-\beta\Delta E}), \quad (1)$$

where $\Delta E = E_i - E_B$ is the difference of potential energies between the quenched replica (E_i) and the baseline replica (E_B), $\beta = 1/kT$, k is Boltzmann’s constant and T is the baseline temperature. After comparison, either the quenched replica (if it is accepted) or the baseline replica (if the quenched replica is rejected) is saved as a sample in the baseline ensemble that will be used for the final statistical analysis of sampled states. The purpose of the quenching stage is to equilibrate all replicas to the same baseline temperature, thus eliminating differences in thermal energy between replicas that otherwise determines

^{a)}Electronic mail: latourr@clemson.edu

the temperature spacing that must be used in a conventional REMD simulation to provide an acceptable degree of exchange between replicas. This feature enables the number of replicas used in a TIGER2 simulation to be independent of the size of the molecular system, allowing users to set the computational cost of the simulation to match the available computational resources at hand.

We note that in the limit of infinite sampling time, the quenched replicas will asymptotically sample from the same distribution as the baseline replica, which would then obviate the application of the Metropolis criterion for exchange. However, this would require that the quenching phase of the cycle (τ_{quench}) greatly exceed the time needed for the full system to equilibrate (τ_{eq}) at the baseline temperature (i.e., $\tau_{\text{quench}} \gg \tau_{\text{eq}}$). Such an approach (involving infinitely slow quenching with no exchange attempts) would be more akin to a simulated annealing approach than a replica-based approach. The intent of replica-based algorithms with exchange, on the other hand, is to make use of the high-temperature sampling without needing to explicitly converge each replica in the baseline temperature ensemble. The TIGER2 method makes use of the fact that the conformational equilibration times are considerably slower than kinetic thermalization times (τ_{therm}). Thus, we essentially require that $\tau_{\text{eq}} \gg \tau_{\text{quench}} \gg \tau_{\text{therm}}$ to assume that the fast-moving portions of the system (e.g., solvent) have rethermalized at the baseline temperature while the slow-moving conformational degrees of freedom (e.g., protein conformation) remain in a configuration sampled from the high-temperature distribution.

The temperature range for a TIGER2 simulation is typically set to enable exploration of all of the relevant phase space of the system. Because the degeneracy of configuration space tends to increase with energy, each high-temperature replica explores a region of configuration space with a relatively narrow range of energies. (This is why many replicas are needed in a REMD simulation.) But because the exchanges are attempted with a *random* quenched state, more states are able to be explored across the configuration space that is sampled by all of the replicas. Thus, even though microscopic reversibility is not strictly preserved, it is maintained in an approximate sense. The closer the collection of replicas comes to providing complete sampling of a relevant subset of configuration space, the closer the overall sampling process comes to providing a properly Boltzmann-weighted distribution of states.

To help illustrate this point, it is helpful to consider two extreme cases. When a TIGER2 simulation is conducted with only one replica, it is equivalent to a MD simulation. Conversely, when a TIGER2 simulation is conducted with a large enough number of replicas to ensure complete coverage of the relevant configuration space, it is equivalent to a Monte Carlo (MC) simulation. In each case, the method is ergodic and generates a Boltzmann-distributed ensemble. The success of the TIGER2 method is based on the idea that the MD baseline is sampled from a Boltzmann distribution, while the quenched replicas act as a reservoir of non-Boltzmann distributed states, at least for the slow degrees of freedom, which are randomly selected and accepted into the baseline MD sequence if the Metropolis criterion is satisfied, thus providing Boltzmann-weighted sampling.

The TIGER2 method has been validated by comparison with REMD simulation results for the folding behavior of relatively small peptides and proteins that could be handled by both methods, including alanine dipeptide (ALAD), (AAQAA)₃ [(Ala-Ala-Gln-Ala-Ala)₃], and chignolin (Gly-Tyr-Asp-Pro-Glu-Thr-Gly-Thr-Trp-Gly) peptides,¹ and protein G² in aqueous solution represented by an implicit solvent model. The TIGER2 results closely matched the REMD results for each of these model systems, thus demonstrating that the TIGER2 algorithm is able to closely approximate Boltzmann statistics.

However, in our recent application of the TIGER2 method in studying aqueous solution of proteins with explicit solvent using the CHARMM/CMAP protein force field,⁵ we found that the energy fluctuations of the solvent prevented the TIGER2 method from properly sampling the conformational behavior of the protein. Similarly, Henriksen *et al.*⁶ recently also found that the TIGER2 method undersampled the low energy states while oversampling the high energy states of alanine dipeptide in explicit water using the AMBER12 force field.^{7,8}

In a biomolecular simulation, explicit solvent is often necessary for the accurate representation of the system. Unfortunately, the number of solvent degrees of freedom required typically greatly outnumbers those needed to represent the solute (e.g., peptide or protein). These extra solvent degrees of freedom not only increase the computational time required per dynamics time step but also increase the statistical sampling errors as well. The variance in the energy, for example, increases by a factor of N when the system size is increased by a factor of N , with explicit water overwhelmingly contributing to the variance in the system's energy.

This additional statistical “energy noise” contributed by the solvent can affect sampling algorithms in two distinct ways: it can increase the sampling time required to reach a desired level of *precision*; and for some sampling methods, such as TIGER2, it can affect the *accuracy* of the sampling as well. To investigate these effects, we conducted a series of increasingly complex simulations with different sampling algorithms, beginning with simple one-dimensional (1-D) models and ending with the sampling of the folding behavior of a small protein with explicitly represented solvation. The results from these simulations demonstrate the solute sampling error that can be caused when using the TIGER2 algorithm in the presence of ‘noisy’ background energy due to a solvent bath and show that the TIGER2A algorithm is able to restore sampling in such systems that closely approximates a Boltzmann-weighted distribution of states.

II. METHODS

A. Simple Metropolis MC simulations with “noisy” solvent bath

To illustrate the intrinsic problem with the TIGER2 algorithm caused by the presence of explicit solvent, consider a Metropolis MC simulation using the following model system, which is designed to represent a solute in a “noisy” solvent bath.

In this simple model system, the solute is represented by a single degree of freedom, x_p , and the bath (solvent) by a single

degree of freedom, x_s . The solute has only two discrete states, $x_p = x_1$ and $x_p = x_2$, with $E(x_1) = 0$ and $E(x_2) = kT$; thus resulting in $\Delta E_p = E(x_2) - E(x_1) = kT$. The solvent degree of freedom is continuous and distributed such that E_s is uniformly distributed between 0 and M . The total system energy is taken to be the sum of the solute and solvent energies, $E = E_p + E_s$, i.e., the solute and bath are not coupled but rather they each independently contribute to the overall energy of the system. We contrast two different approaches to performing MC-based sampling for this system.

In the first method, the configurational state of the system is given by (x_p, x_s) , the joint state of both the solute and the solvent, which has some associated total energy $E = E_p + E_s$. A MC step involves the generation of a new configurational state, (x'_p, x'_s) , for both the solute and bath, with a total energy of E' , and that the new state is accepted with probability $p = \min(1, e^{-(E'-E)/kT})$. Accordingly, the new sampled state is automatically accepted if E' is lower than E , or if E' is greater than E , the new state is accepted only if the Metropolis criterion is satisfied. If the new state is rejected, the prior state (x_p, x_s) is accepted and compared again with the next newly sampled state. This algorithm, of course, represents the standard Metropolis sampling procedure and generates a Boltzmann distribution, and the only effect of the solvent noise is to hinder the speed of convergence towards the Boltzmann average.

For the solute degree of freedom, the expected population of the lower-energy state (x_1) is

$$p_1 = \frac{e^{-E_1/kT}}{e^{-E_1/kT} + e^{-E_2/kT}} = \frac{1}{1 + e^{-1}} \approx 0.73. \quad (2)$$

Figure 1 shows the convergence of p_1 to this limit with increased sampling for several different values of the noise magnitude M ranging from 0 to 500 kT . For large values of M , the acceptance of any individual move is driven primarily by the value of the bath energy. But because M is randomly sampled over a uniform distribution, the average difference in the solvent energy between each of the selected new states and the previously accepted state is zero (i.e., $\overline{E'_s - E_s} = 0$), so the

FIG. 1. The effect of bath energy noise on the convergence of Monte Carlo sampling towards the Boltzmann distribution for a two-state model system for magnitudes M of the bath noise as provided in the legend. The solid line indicates the Boltzmann result.

equilibrium average is determined only by the solute states, and the Boltzmann distribution is recovered after sufficient sampling. As shown in Figure 1, at large noise levels, it simply takes a lot more sampling to reduce the noise from the bath to a level small enough to make meaningful measurements of the signal (i.e., ΔE_p) from the solute, with the probability of sampling state x_1 converging to about 0.73.

Now, consider an alternative MC-like sampling algorithm, inspired by replica-based simulation methods. In this method, each replica is typically propagated using MD for some period of time, after which it may be replaced by a different replica, through an exchange mechanism. The expectation is that the fast-moving bath (solvent) degrees of freedom update frequently within the MD cycle, while changes to the slow-moving solute (e.g., a protein), collective degrees of freedom often come from exchanges. Consequently, we can consider the system to begin in some state (x_p, x_s) with total energy $E = E_p + E_s$. After the MD cycle, the state is (x_p, x'_s) , where we have assumed that the fast bath modes have been resampled but the slow solute modes have not. Thus, the energy is $E' = E_p + E'_s$. Next, a new configuration (x''_p, x''_s) is generated, with energy $E'' = E''_p + E''_s$, corresponding to a state that might be sampled from another replica. This move is then accepted with probability $p = \min(1, e^{-(E''-E')/kT})$. If the move is accepted, the state (x''_p, x''_s) contributes to the ensemble average; if it is rejected, then the state (x_p, x'_s) contributes. Note that the only difference between this method and the prior MC method is that when a move is rejected and a state contributes twice to the ensemble, the bath coordinate x'_s is resampled in this method, with its energy contribution ranging between 0 and M . Thus, for the case where $M \gg kT$, the solvent energy will have a dominant effect in determining whether the new state is accepted or not, with little sensitivity to the difference in energy between state x_p and x''_p . This approach is flawed and generates non-Boltzmann sampling but might at first glance appear to be a reasonable approach for replica-based sampling, in which the bath motions are very fast and contribute large energies compared to solute motions and are resampled frequently within one MD cycle.

Figure 2 illustrates the convergence towards equilibrium for this sampling method on the two-state system for various magnitudes of bath noise. When $M = 0$, the occupation probabilities converge to the Boltzmann distribution. But for nonzero noise, they converge towards a non-Boltzmann limit, with the limiting value of p_1 being closer to 0.5 at increasing levels of M . Larger bath noise results not only in slower convergence but also a systematic bias away from the Boltzmann equilibrium.

The reason for this discrepancy is that the sampling algorithm violates detailed balance. When the noise is resampled, the system has made a transition from state (x_p, x_s) to state (x_p, x'_s) . Despite the fact that x'_s was drawn from the appropriate probability distribution, and the bath modes will thus be distributed correctly, the ratio of this forward move and backward move is not the ratio of the Boltzmann probabilities of the states, and detailed balance is not obeyed. When the magnitude of the noise is large compared to ΔE_p , the acceptance of the move will be driven by the energy difference $E''_s - E'_s$ between two randomly sampled bath states, rather than the energy difference $E''_p - E_s$ between a new (randomly sampled)

FIG. 2. The effect of bath energy noise on convergence of the two-state model for a sampling algorithm in which the bath noise is resampled before every move attempt. Magnitudes M of the bath noise are provided in the legend. The solid line indicates the Boltzmann result.

state and the prior accepted state. The energy difference $E_s'' - E_s'$ between two randomly sampled states is symmetrically distributed about zero, for any distribution of bath noise, and thus, the probability of accepting any move will approach 50% when $|E_s'' - E_s| \gg kT$. This explains why the occupation probability p_1 approaches 0.5 for large M in Figure 2.

The flaws in the latter method are fairly obvious when presented in the context of this simple one-dimensional model with applied noise. The purpose of discussing it at length is not, of course, to introduce a flawed method for performing conformational sampling. Instead, it is to provide some illumination of approaches used when performing conformational sampling in more realistic systems where the bath noise is resampled very quickly compared to slow solute motions. What this example shows is that if the moves (or exchanges) are attempted between a newly sampled state and a prior state *with resampled solvent coordinates*, then detailed balance is not obeyed, and the consequence is that the distribution will be biased away from Boltzmann towards a flat distribution that is largely insensitive to the energy of the solute. The previous TIGER2 implementation suffered from this flaw when used with explicitly represented solvent. The current TIGER2A algorithm is designed to remedy this deficiency.

To understand how this is relevant to a TIGER2 sampling process, let the total potential energy of each replica, after quenching and equilibration at the baseline temperature level in a TIGER2 simulation, be represented by

$$E_i = E_{i,pp} + E_{i,ps} + E_{i,ss}, \quad (3)$$

where E_i is the total potential energy of the i th replica and E_{pp} , E_{ps} , and E_{ss} are the protein–protein, protein–solvent, and solvent–solvent interactions, respectively. With this decomposition, the ΔE term in Eq. (3) can be expressed as

$$\Delta E = \Delta E_{pp} + \Delta E_{ps} + \Delta E_{ss}, \quad (4)$$

where ΔE_{pp} , ΔE_{ps} , and ΔE_{ss} measure the differences of protein–protein, protein–solvent, and solvent–solvent interactions between the quenched and the baseline replicas.

In a system containing a large amount of solvent, the magnitude of ΔE_{ss} will in most cases be much larger than that of

ΔE_{pp} or ΔE_{ps} , so the MC sampling process will be dominated by the solvent energy difference between the quenched and baseline replicas. If the quenched replica has a particularly favorable solvent–solvent energy at the instant when it is sampled, the move is likely to be accepted, even when the protein conformation has increased in energy. Likewise, if the quenched replica has an unfavorable solvent–solvent energy compared to the baseline replica, the move is likely to be rejected, even when the protein conformation is lower in energy. And crucially, as described above, the resampling of the solvent state over the course of the next cycle prevents these solvent configurations from being accepted with Boltzmann probabilities, leading to the under-population of the solute’s low-energy states and the over-population of its high-energy states.

In the present work, we present a new version of the TIGER2 method specifically for simulations of systems containing explicit solvent. The new method is named “TIGER2 with solvent energy averaging” (TIGER2A) and is able to closely approximate a Boltzmann ensemble at the baseline temperature in the presence of explicit solvent, thus solving the sampling problem encountered with TIGER2. In this method, averaging is used to reduce the solvent energy fluctuation to a level below the changes of solute–solute and solute–solvent interactions so that the exchanges between a quenched replica and the baseline replica are determined primarily by the energetics of the solute, rather than noisy solvent fluctuations. The TIGER2A method is introduced in detail in Sec. II B and demonstrated for systems of increasing complexity in Sec. III.

B. The TIGER2A algorithm

In a TIGER2A simulation for a system containing a solute in explicitly represented solvent, the sampling procedure consists of a series of cycles with each cycle containing: (1) rapid heating from a baseline temperature (T_b) to the replica temperature (T_i) by rescaling the momenta of the atoms within the replica by a factor of $\sqrt{T_i/T_b}$ and thermally equilibrating for a time t_1 , (2) MD sampling at constant temperature (T_i) for a time t_2 , (3) rapid quenching of replicas sampled at elevated temperatures back down to T_b by rescaling the momenta by a factor of $\sqrt{T_b/T_i}$ followed by thermal equilibration and sampling for a time t_3 , (4) further equilibration of the solvent degrees of freedom, while holding the solute in a fixed conformation for a time t_4 in which the energy is averaged, (5) comparison between the averaged potential energies of the baseline and quenched replicas using the Metropolis criterion, and (6) global replica reassignment to the temperature levels for the subsequent sampling cycle.

This TIGER2A algorithm differs from the previous TIGER2 algorithm^{1,2} in stages (4) and (5). TIGER2 does not include stage (4) and the Metropolis comparisons in stage (5) are performed with the instantaneous replica energies (Eq. (4)) from the end of stage (3). TIGER2A, on the other hand, adopts an equilibration stage (4) in which the conformation of the solute is fixed and the solvent energy is averaged, and the Metropolis comparisons in stage (5) are performed with energies

$$E_i = E_{i,pp} + \langle E_{i,ps} \rangle + \langle E_{i,ss} \rangle, \quad (5)$$

where the averaging is performed over all solvent configurations generated in stage (4). While ideally the solute dynamics should occur under the potential of mean force (PMF), rather than the averaged solvent energy, the calculation of a true PMF would substantially complicate the quenching phase of the algorithm. As we demonstrate in Sec. III, the use of the average of the solvent energy greatly reduces the sampling error encountered with TIGER2 and enables a Boltzmann-weighted ensemble of states to be closely approximated by the TIGER2A sampling process for simulations that use explicit solvent.

After sufficient averaging at the baseline temperature, the contributions from the solvent fluctuations can be made arbitrarily small, so that the exchanges are driven primarily by the energetic contributions arising from solute (e.g., protein) conformational differences. The number of steps for the equilibration in stage (4) is thus determined empirically based on the condition that

$$|\langle E_{i,ss} \rangle - \langle E_{b,ss} \rangle| \leq |(\langle E_{i,pp} \rangle + \langle E_{i,ps} \rangle) - (E_{b,pp} + \langle E_{b,ps} \rangle)|, \quad (6)$$

by which we mean that the difference in solvent-solvent energy between quenched and baseline replicas is less than the difference in interaction energies involving the solute.

In stage (5), the baseline replica is first copied to a pool of accepted states and then used as a reference state to compare with each of the other replicas. The quenched replicas serve as a reservoir of states⁹ and in turn, each of them in increasing order of the temperature that they were quenched from is compared against the reference state and selected with the Metropolis criterion in Eq. (1). If the new state is selected, it is copied to the pool of accepted states and then becomes the new reference state for the next comparison. If not, then the prior reference state remains for the next comparison. A replica is not copied to the baseline ensemble multiple times per TIGER2A cycle. After all replicas have been considered as potential ensemble members, the replicas are reassigned to the temperature levels in stage (6). In this stage, the last reference state is assigned to the baseline temperature and the rest of states are then reassigned to the higher temperature levels according to their potential energies; i.e., a higher potential energy state is assigned to a higher temperature level. Note that stage (5) of TIGER2A (Metropolis comparison) differs from stage (4) of TIGER2. In this step, TIGER2 conducts only one Metropolis comparison and saves a maximum of one sample to the accepted-state ensemble per cycle, whereas TIGER2A conducts $N_r - 1$ Metropolis comparisons and always saves at least one sample to the ensemble of accepted states. Therefore, a TIGER2A cycle is able to generate more states than a TIGER2 cycle, which compensates for the longer cycle time of TIGER2A from the addition of the solvent-energy averaging stage of the cycle.

C. Model systems for assessment of TIGER2A performance

Four distinctly different model systems were used to evaluate the performance of the TIGER2A algorithm in comparison with TIGER2 and to demonstrate the ability of TIGER2A

to provide a closer approximation to Boltzmann sampling for each system: (1) a simple 1-D analytical model, (2) a single ALAD molecule in implicit water represented by the Generalized Born model with a simple switching (GBSW) function,^{10,11} (3) the small peptide (AAQAA)₃ in implicit water represented by the GBSW model, and (4) hen egg-white lysozyme (HEWL) protein in explicit water represented by the TIP3P water model.^{12,13} For model systems 1-3, solvent energy fluctuations were simulated for each step of the MD simulation by randomly selecting a value for background energy noise from a Gaussian distribution of energies obtained from a separate set of explicit water simulations.

The simulations for the 1-D model were conducted with an in-house developed FORTRAN program and the simulations for systems (2) through (4) were carried out with CHARMM molecular modeling software^{14,15} using the CHARMM22 force field with CMAP correction.⁵ The dynamics for system (1) were performed with a velocity Verlet integrator (VV2) and Langevin thermostat. The calculations for systems (2) through (4) were conducted with the modified VV2^{16,17} and a Nosé-Hoover thermostat^{18,19} with a time step of 2 fs. Bonds involving hydrogen atoms were constrained using the SHAKE²⁰ algorithm. The non-bonded van der Waals and electrostatic interactions were treated with a real space cutoff of 20 Å for systems 2 and 3 and 9 Å for system 4 respectively.

1. 1-D analytical model

As a 1-D model system to be used in an investigation of the sampling methods considered here, we define a total system potential energy to be composed of contributions from a single solute degree of freedom, φ , and one uncoupled solvent degree of freedom, q , so that

$$E(\varphi, q) = E_p(\varphi) + E_s(q). \quad (7)$$

The one-dimensional solute energy [$E_p(\varphi)$] and the corresponding force [$F_p(\varphi)$] acting on the solute are defined by

$$E_p(\varphi) = \frac{1}{2} \sum_{i=1}^n A_i [1 + (-1)^{i-1} \cos(i\varphi)], \quad (8)$$

$$F_p(\varphi) = -\frac{dE_s(\varphi)}{d\varphi} = -\frac{1}{2} \sum_{i=1}^n A_i [(-1)^i i \sin(i\varphi)], \quad (9)$$

where the A_i are a force constants that define the potential. The series of A_i is defined by

$$\begin{aligned} A_1 &= \lambda, \\ A_2 &= \frac{1}{2} \left(\frac{1}{2} \lambda + \sigma + \sqrt{(\lambda + \sigma)\sigma} \right), \\ A_i &= \frac{1}{2} \left(\frac{1}{2i-3} \lambda + \sigma + \sqrt{(\lambda + \sigma)\sigma} \right), \quad \text{for } i \geq 3. \end{aligned} \quad (10)$$

The purpose of this model system is to provide a 1-D potential for which the Boltzmann distribution can be determined exactly, but with a range of different barrier heights, and for which the convergence toward the Boltzmann distribution can be easily adjusted. An example of the potential energy of a five-state model, i.e., $i = 5$, is shown by Figure 3.

The solvent is represented by the aggregated effect of N independent solvent degrees of freedom, each moving in

FIG. 3. The analytical energy landscape (A) and the probability distribution (B) given by the five-fold 1-D model.

a harmonic well with spring constant k_{sp} . For one of the N solvent degrees of freedom, the potential energy, $E_s(q)$, and the corresponding force, $F_s(q)$, are defined by

$$E_s(q) = \frac{1}{2}k_{sp}q^2, \quad (11)$$

$$F_s(q) = -\frac{dE_b(q)}{dq} = -k_{sp}q. \quad (12)$$

When simulated in the canonical ensemble, the average solvent energy is

$$\langle E_s \rangle = \frac{1}{2}NkT. \quad (13)$$

The solvent energy can be made arbitrarily large or small, with respect to the solute energy, by changing N . Note that in this simple model, the $N + 1$ degrees of freedom (i.e., solvent + solute) are modeled with the computational expense of only explicitly simulating two coordinates, φ and q .

Eq. (8) with $n = 5$ was selected to represent a five-term potential characterized by a rugged energy surface with multiple potential wells for the solute. The force constants A_1 – A_5 given by Eq. (10) were established with $\lambda = kT_0$ and $\sigma = 4kT_0$, where kT_0 is an arbitrary constant that defines the energy scale of the potential, resulting in the potential plotted in Figure 3. Five replicas were used in both TIGER2 and TIGER2A simulations with uniformly distributed temperature levels $T = kT_0$, $3kT_0$, $5kT_0$, $7kT_0$, and $9kT_0$.

For TIGER2, three simulations labeled as T-1D-1, T-1D-2, and T-1D-3 with bath sizes of $N = 0$ (no solvent), 2×10^2 , and 2×10^3 were used to examine the effect of solvent fluctuations on the accuracy of the method. For these systems, the standard deviations of the solvent energy are 0, 14.762, and 146.746 kT_0 at the baseline temperature, respectively. These values for the thermal fluctuation of the solvent can be compared to the energy differences of about 12 kT_0 between the lowest and highest energies of the solute. The TIGER2 simulations were conducted for 2×10^5 cycles to obtain convergence. Each TIGER2 cycle included 10^3 steps of heating, 10^3 steps of sampling, and 10^2 steps of quenching. For TIGER2A, three simulations labeled as TA-1D-1, TA-1D-2, and TA-1D-3 were conducted with $N = 2 \times 10^2$ (similar to T-1D-2), but with 10^4 , 10^3 , or 10^2 steps of solvent-energy averaging applied to examine the effects of different levels of reduction in the statistical noise due to the energy fluctuations of the simulated solvent. The TIGER2A simulations were conducted for 3×10^6 cycles to obtain convergence.

The convergence of the simulation results towards the analytical result of this 1-D model was measured by the L^2 norm, or root mean square deviation (RMSD) of the sampled probability distribution, $P(\varphi)$, from the corresponding analytical distribution, $P_{eq}(\varphi)$, at the baseline temperature,

$$RMSD = \left(\frac{1}{2\pi} \int_0^{2\pi} [P(\phi) - P_{eq}(\phi)]^2 d\phi \right)^{1/2}. \quad (14)$$

2. ALAD

This simulated system used alanine dipeptide (Ace-Ala-Nme) as the solute. Solvation effects were represented by the GBSW implicit model.^{10,11} Default GBSW parameters were used with a 0.6 Å smoothing length along with 50 Lebedev angular integration points and 24 radial integration points up to 20 Å for each atom. The nonpolar solvation energy was estimated from the solvent-exposed surface area using a phenomenological surface tension coefficient of 0.005 kcal mol⁻¹ Å². We used simulated background noise to mimic the solvent-solvent interactions by randomly generating solvent energy contributions based on a Gaussian distribution function, i.e.,

$$E_{i,ss} = \langle E \rangle + w x_i, \quad (15)$$

where x_i is a Gaussian pseudo-random number with zero mean and unit variance for the i th replica generated via the Box-Muller transform.²¹ $\langle E \rangle$ and w are the mean and standard deviation of the energy distribution, which we assume to be represented well by a system of explicit water. For a periodic box with 5003 TIP3P waters (WAT1), we obtained $\langle E \rangle = -48\,897$ kcal/mol and $w = 201$ kcal/mol. For another box containing 25 020 TIP3P waters (WAT2), we obtained $\langle E \rangle = -244\,252$ kcal/mol and $w = 441$ kcal/mol. These values were then used in Eq. (15) to generate synthetic noise signals in both the TIGER2 and TIGER2A simulations to represent the energy fluctuation in the system caused by a water bath.

The REMD simulation was conducted with 24 replicas with a GBSW representation of solvation effects, but without the synthetic background noise. The 24 temperature levels for the REMD simulation were exponentially distributed between 300 and 608 K. Two TIGER2 simulations labeled as T-AD-1 (using synthetic WAT1) and T-AD-2 (using synthetic WAT2) were conducted using twelve replicas with temperature levels uniformly distributed between 300 K and 608 K. Two TIGER2A simulations labeled as TA-AD-1 (using synthetic WAT1) and TA-AD-2 (using synthetic WAT2) were

conducted using a similar protocol as the TIGER2 simulations for performance comparison to TIGER2. The starting structure of each replica was randomly selected from different ϕ - ψ positions in the Ramachandran plot of ALAD. Each replica was conducted for 4×10^4 cycles in total. In the TIGER2 simulations, each cycle included 1.0 ps of heating/equilibration, 3.0 ps of sampling, and 2.0 ps of quenching/equilibration, with solvent energy added at the end of the cycle based on the value selected from the Gaussian distribution of solvent energy per Eq. (15). The same cycle design was used for the TIGER2A simulations except instead of the single value of energy representing solvent energy fluctuations, an average value of the solvation energy was added, which was obtained as the average of 10^4 independent samples from the Gaussian distribution of water-box energy using Eq. (15).

The simulations were performed and comparisons were then made of the sampled probabilities of the various ϕ - ψ regions of the Ramachandran plot provided for each case compared to the equilibrated REMD results.

3. Helical peptide (AAQAA)₃

The (Ala-Ala-Gln-Ala-Ala)₃ [or (AAQAA)₃] peptide is an alanine-based synthetic model peptide containing 15 amino acid residues that fold into an α -helical structure. In the simulations conducted in this work, the peptide was capped with the N-terminal acetyl (Ace) and C-terminal methylamino (Nme) capping groups. A similar GBSW^{10,11} protocol and the method to generate synthetic solvent noise was used for these simulations as was used for the ALAD simulations. All simulations were conducted using twelve replicas with temperature levels uniformly distributed between 270 K and 600 K. To obtain initial states for the twelve replicas, we selected twelve configurations from a TIGER2 simulation of (AAQAA)₃ with implicit solvent represented by the GBSW model (i.e., without background solvent energy noise fluctuations). These structures contained one completely folded state, eight partially folded states and three random coil states. The energies of the twelve states varied from -52.543 to 12.781 kcal/mol (note the small variation in solute energies, relative to the solvent fluctuations with standard deviations of 201 and 441 kcal/mol for the small and large water boxes, respectively).

Each simulation was conducted for 1.8×10^4 cycles in total. In the TIGER2 simulations, each cycle included 1.0 ps of heating/equilibration, 3.0 ps of sampling, and 2.0 ps of quenching/equilibration. As with the ALAD simulations, to represent the effect of background solvent energy “noise” on the sampling process, a single solvent energy value selected from the Gaussian distribution of solvent energy per Eq. (15) was added to the solute’s energy prior to the Metropolis comparison for the TIGER2 simulations. For the TIGER2A simulations, an averaged value of solvation energy from 10^4 randomly selected solvent energy noise values was added prior to each Metropolis comparison.

These simulations were performed and comparisons were then made of the percent helicity of each of the amino acids in this peptide for each sampling condition compared to previously published REMD results.¹

4. HEWL simulation

HEWL is a relatively small (129-residue) enzyme that catalyzes the hydrolysis of specific kinds of polysaccharides comprising the cell walls of bacteria.^{22,23} In preliminary conventional MD simulations, HEWL was determined to retain its native state structure when simulated in explicit water at 300 K. In the present work, the TIGER2 and TIGER2A methods were used to sample the conformation of HEWL in aqueous solution, with the expectation that the sampling algorithm should retain HEWL’s native-state structure (as represented by its percent helicity) at 300 K, with an increasing degree of conformational unfolding (i.e., loss in helical structure) from the elevated temperature replicas.

The modeled system contains a HEWL protein (PDB entry 1GXV²⁴) in a water box with dimensions $85 \times 85 \times 85 \text{ \AA}^3$. Periodic boundary conditions were applied in all dimensions. Both TIGER2 and TIGER2A simulations used twenty-four replicas with temperature levels being uniformly distributed between 300 and 691 K. To generate the initial replica structures, lysozyme coordinates were randomly selected from a MD trajectory at 300 K. The TIGER2 sampling cycle contained 2 ps of heating and equilibration followed by 46 ps sampling at the elevated temperature, and then 14 ps quenching and equilibration back to the baseline temperature, with the final potential energy of each replica being used for the Metropolis comparison at the end of the cycle. For TIGER2A, the same times were used for the heating, sampling, and quenching stages of the cycle, followed with a 50-ps additional stage in which the protein was kept fixed in position while the water was allowed to move, with the energy of the system then averaged over this 50 ps timeframe to provide the final energy values for each replica that were then used for the Metropolis comparisons.

The resulting sampled distribution of states for each sampling method was then evaluated and compared by monitoring the RMSD of HEWL compared to its native-state structure and its percent helicity vs. sampling time at both the baseline temperature (300 K) and at the maximum temperature (691 K).

III. RESULTS AND DISCUSSION

A. 1-D analytical model

The conformational energy of the five-term 1-D model defined by Eqs. (8) and (10) is shown in Figure 3(A) and the corresponding probability distribution in Figure 3(B). The potential is periodic on $[0, 2\pi]$, and symmetric about $\varphi = \pi$,

TABLE I. Definition of four characteristic regions for the five-term 1D potential, with corresponding local energy minima and Boltzmann population of the region at $T = T_0$.

Region	Domain range (rad)	$E_{\min}(\varphi)/kT$	$P_i(\varphi)$
1	$[0, 0.807)$ and $[5.473, 2\pi)$	7.8160	0.040
2	$[0.807, 1.850)$ and $[4.430, 5.473)$	5.5182	0.275
3	$[1.850, 2.660)$ and $[3.620, 4.430)$	7.1795	0.085
4	$[2.66, 3.620)$	4.3365	0.600

FIG. 4. (A) Convergence of three TIGER2 simulations, T-1D-1, T-1D-2, and T-1D-3, towards the Boltzmann distribution with sampling time as measured by the root mean square deviation (*RMSD*) of calculated probability from the corresponding theoretical result. (B) Comparison between the theoretical and calculated probability distributions. For TIGER2 simulations, the distributions were calculated based on the last 10^5 cycles.

so it contains 4 symmetrically distinct minima. We divided the domain of φ into four characteristic population regions (Table I), with boundaries at the local maxima of $V(\varphi)$. The Boltzmann populations of these four regions at $T = T_0$ are given in Table I.

The convergence of the three TIGER2 simulations, i.e., T-1D-1, T-1D-2, and T-1D-3, towards the Boltzmann distribution is shown in Figure 4. Figure 4(A) shows the evolution of the *RMSD* convergence measure [Eq. (14)] with increasing sampling time, while Figure 4(B) compares the probability distributions calculated from T-1D-1 and T-1D-3 with the analytical probability distribution. The profile obtained from simulation T-1D-2 is not presented to keep the figure legible. The integrated population densities in each of the four characteristic regions are given in Table II.

The results from Figure 4 and Table II illustrate that energy fluctuations from the solvent can cause the TIGER2 sampling to deviate from the canonical ensemble. A comparison of Tables I and II shows that the T-1D-1 simulation, which is not influenced by solvent fluctuation, samples the different potential wells with probabilities in close agreement with the theoretical results, and Figure 4(A) shows that it converges close to this theoretical limit with additional sampling time. Simulations T-1D-2 and T-1D-3, on the other hand, display significant differences from the analytical result. The solvent fluctuations cause these simulations to oversample the higher energy regions and undersample the low energy well. This behavior is analogous to what was shown in the two-state noisy Monte Carlo example presented in Figure 2, where the acceptance of new states is dominated by the value of the

solvent energy “noise,” and random resampling of the solvent state before the next cycle represents a violation of detailed balance. Figure 4(A) illustrates that this is not merely a statistical sampling error: the simulations converge in the limit of long sampling time to a probability distribution that is different from the Boltzmann distribution, with errors that are larger for larger bath sizes.

For the three TIGER2A simulations (TA-1D-1, TA-1D-2 and TA-1D-3), Figure 5(A) shows the convergence of these simulations and 5(B) compares the calculated and analytical probability profiles (again omitting TA-1D-2, for clarity). The integrated population densities of the four characteristic regions are presented in Table II. When averaging for only a small number of steps, as in TA-1D-3, the situation is only slightly improved to that for T-1D-2 (TIGER2 simulation with same N), and the high-energy states are still oversampled at the expense of the lower-energy states. But as the averaging time increases, the magnitude of this oversampling error decreases, as the fluctuation of the averaged solvent energy is reduced. Through simulations TA-1D-3 to TA-1D-1, SD_{solvent} is reduced from $10.836 kT_0$ to $1.602 kT_0$ (the standard deviation of the solute energy is about $1.200 kT_0$, for comparison), and the *RMSD* between the sampled and Boltzmann probability distributions decreases from about 0.100 to 0.027, as shown in Figure 5(A).

B. Alanine dipeptide

For the alanine dipeptide in aqueous solution, the main features of the backbone dihedral angle ϕ - ψ distribution are

TABLE II. Effect of solvent noise and averaging on the population probabilities of the four characteristic regions at baseline temperature for TIGER2 (T-1D-1 to 3) and TIGER2A (TA-1D-1 to 3) simulations for the 5-fold 1D model.

Simulation	Averaging steps	$SD_{\text{solvent}} (kT_0)^a$	$P_1(\varphi)$	$P_2(\varphi)$	$P_3(\varphi)$	$P_4(\varphi)$
T-1D-1	0	0	0.034 ± 0.001	0.268 ± 0.001	0.092 ± 0.001	0.606 ± 0.001
T-1D-2	0	14.762	0.118 ± 0.002	0.294 ± 0.002	0.104 ± 0.001	0.484 ± 0.002
T-1D-3	0	146.746	0.066 ± 0.002	0.465 ± 0.003	0.065 ± 0.002	0.404 ± 0.005
TA-1D-1 ^b	10 000	1.602	0.066 ± 0.001	0.263 ± 0.001	0.079 ± 0.001	0.592 ± 0.001
TA-1D-2 ^b	1 000	4.124	0.101 ± 0.001	0.284 ± 0.002	0.094 ± 0.001	0.521 ± 0.002
TA-1D-3 ^b	100	10.836	0.118 ± 0.003	0.296 ± 0.002	0.098 ± 0.001	0.488 ± 0.003

^aStandard deviation of the solvent energy in kT_0 . For the T-1D-2,3 simulations, *SD* represents the standard deviation of the solvent energy. For the TA-1D-1,2,3 simulations, *SD* represents the standard deviation of the averaged energy values after the designated number of averaging steps were applied. For comparison, the standard deviation of the solute energy fluctuation in vacuum (no solvent fluctuation) is $1.2 kT_0$.

^bSimulation was conducted with a bath size $N = 200$ in Eq. (13).

FIG. 5. (A) Convergence of three TIGER2A simulations, TA-1D-1, TA-1D-2, and TA-1D-3, towards the Boltzmann distribution with sampling time as measured by the root mean square deviation (*RMSD*) of calculated probability from the corresponding theoretical result. (B) Comparison between the theoretical and calculated probability distributions. For TIGER2A simulations, the distributions were calculated based on the last 10^5 cycles.

six regions corresponding to conformations designated $C7_{eq}$, $C5$, $C7_{ax}$, β_2 , α_R , and α_L , with local energy minimum centered around $-82^\circ/75^\circ$, $-160^\circ/64^\circ$, $78^\circ/-56^\circ$, $-138^\circ/23^\circ$, $-60^\circ/-45^\circ$, and $64^\circ/35^\circ$, respectively, as predicted with the CHARMM/CMAP force field.⁵ Using this as our model system, we divided the entire ϕ - ψ map into four characteristic regions corresponding to $C7_{eq}/C5$, α_R/β_2 , α_L , and $C7_{ax}$ conformers. The boundaries between these regions of the ϕ - ψ map are indicated in Fig. 5 of Ref. 25. Based on 10^5 configurations sampled from REMD, TIGER2, and TIGER2A simulations, we calculated the population ratios of conformers located in these four regions and then quantitatively contrasted the performance of the TIGER2 and TIGER2A methods with REMD.

The calculated mean values and standard deviations of the population ratios of conformers located in the $C7_{eq}/C5$, α_R/β_2 , α_L , and $C7_{ax}$ regions are presented in Table III and labeled as P_A , P_B , P_C , and P_D correspondingly. Figure 6 displays the P_A and P_B populations of alanine dipeptide calculated by block averages over every 10^4 samples in REMD ($\blacksquare$), T-AD-1 ($\blacktriangle$), T-AD-2 ($\blacktriangledown$), TA-AD-1 ($\triangle$), and TA-AD-2 (∇) simulations. In all simulations, the block averaged population ratios plateaued to equilibrium values. The results shown in Table III indicate that the synthetic solvent noise in the two TIGER2 simulations, T-AD-1 and T-AD-2, tend to generally over-sample the high energy conformers (i.e., $C7_{eq}/C5$, α_L , and $C7_{ax}$) and under-sample the low energy conformers (i.e., α_R/β_2), with these calculation errors being clearly evident when compared to the REMD results as shown in Figure 6. As shown in Figure 6, the two TIGER2A simulations, TA-AD-1 and TA-AD-2, effectively reduce the influence of solvent fluctuation and provide much closer agreement with the REMD results.

C. (AAQAA)₃

The model peptide (AAQAA)₃ is known to fold into a stable α -helix in aqueous solution with an α -helical content of about 50% at 274 K.²⁶ We performed two different TIGER2 simulations (T-MP-1 and T-MP-2) and two TIGER2A simulations (TA-MP-1 and TA-MP-2) for the (AAQAA)₃ peptide, with different levels of synthetic solvent energy noise. The SD_{wat} was 230.866 kcal/mol in the T-MP-1 and TA-MP-1 simulations, and 100.131 kcal/mol in the T-MP-2 and TA-MP-2 simulations. For comparison, the standard deviation of the energy of (AAQAA)₃ in implicit solvent represented by the GBSW model at 270 K is 9.621 kcal/mol. The TIGER2A simulations used 10^4 steps of solvent averaging, which reduced SD_{wat} to 2.232 and 1.009 kcal/mol in the TA-MP-1 and TA-MP-2 calculations, respectively.

In Figure 7, the accuracy of the four simulations is compared with the corresponding REMD results,¹ using the fractional helicity of each residue of the peptide at 270 K, averaged over the last 10^4 cycles of each simulation. The difference between TIGER2 and TIGER2A is particularly clear in this system. The TIGER2A and REMD simulations resulted in similar residue helicities at each residue, whereas the fractional residue helicity calculated by the two TIGER2 simulations were much lower than the corresponding REMD results and close to 0% helicity.

The average helicity for the entire peptide as calculated by REMD, TA-MP-1, and TA-MP-2 is 0.623 ± 0.031 , 0.582 ± 0.030 , and 0.593 ± 0.032 , respectively. These values agree reasonably well, thus demonstrating the ability of TIGER2A to approximate Boltzmann-weighted sampling for this model system in the presence of background energy noise. The average peptide helicity calculated by T-MP-1 and T-MP-2 is

TABLE III. Occupation probabilities of the four characteristic conformers of alanine dipeptide at baseline temperature obtained from REMD, TIGER2 (T-AD-1 and 2), and TIGER2A (TA-AD-1 and 2) simulations.

Simulation	$P_A (\phi, \psi)$	$P_B (\phi, \psi)$	$P_C (\phi, \psi)$	$P_D (\phi, \psi)$
REMD	0.1667 ± 0.0320	0.8292 ± 0.0315	0.0005 ± 0.0013	0.0036 ± 0.0033
T-AD-1	0.2803 ± 0.0207	0.6800 ± 0.0196	0.0053 ± 0.0015	0.0344 ± 0.0060
T-AD-2	0.2813 ± 0.0147	0.6831 ± 0.0147	0.0061 ± 0.0027	0.0295 ± 0.0087
TA-AD-1	0.2073 ± 0.0169	0.7829 ± 0.0183	0.0006 ± 0.0006	0.0092 ± 0.0041
TA-AD-2	0.2142 ± 0.0111	0.7780 ± 0.0104	0.0004 ± 0.0003	0.0074 ± 0.0021

FIG. 6. The P_A and P_B populations of alanine dipeptide calculated by block averages over every 10^4 samples in REMD (—■—), T-AD-1 (—▲—), T-AD-2 (—▼—), TA-AD-1 (—△—), and TA-AD-2 (—▽—) simulations.

both only 0.022 ± 0.010 , indicating that the TIGER2 method is oversampling the (high-energy) unfolded states due to the effects of the solvent noise and is not able to accurately predict the folding behavior of the (AAQAA)₃ peptide.

D. HEWL

The HEWL protein contains two structural domains: (1) the α -domain, consisting of four α -helices and a C-terminal 3_{10} helix and (2) the β -domain, consisting of a triple-stranded antiparallel β -sheet, a 3_{10} helix, and a long loop. The two domains are linked by a short double-stranded antiparallel β -sheet through disulfide bridges.

Starting from the crystal structure of the HEWL protein, 10 ns TIGER2 and 25 ns TIGER2A simulations were performed with temperature levels uniformly distributed from 300 to 691 K. (The duration of a TIGER2 or TIGER2A refers to the total time accumulated by the baseline replica.) The high temperature of 691 K was chosen after initial, exploratory simulations designed to identify the temperature required to fully unfold the protein within a practical simulation timeframe (taken as 10 ns for the sake of our simulations).

The C_α RMSD for the sampled HEWL structures compared with the native-state crystalline structure (PDB ID 1GXV²⁴) for the 300 K and 691 K quenched replicas is plotted

in Figures 8(A) and 8(B) for the TIGER2 and TIGER2A simulations, respectively. Significant differences between the two simulation methods are clearly evident. In the TIGER2 profiles shown in Figure 8(A), both the 300 K and 691 K replicas include a large number of unfolded states (with large C_α RMSD) and a somewhat smaller number of folded states (small C_α RMSD), rather than sampling primarily folded structures at 300 K and unfolded structures at 691 K, as appropriate for Boltzmann sampling of this system. This is again caused by the thermal fluctuation of solvent in the TIGER2 sampling procedure. For example, a low-energy folded state of the HEWL protein at 300 K that happened to be sampled during a high-energy fluctuation of the water energy had a low probability of remaining in the baseline ensemble when this baseline replica was compared to a quenched replica from 691 K that happened to have much lower solvent energy, even though it may have had a much higher-energy unfolded state for the HEWL protein. In other words, the exchanges were driven primarily by the large fluctuations of the solvent energy of the system, which were much larger than the energy difference between folded and unfolded states of the protein.

As seen in Figure 8(B), The TIGER2A simulations perform much better. Figure 8(B) shows that the C_α RMSD of the baseline replica fluctuates around 3.2 Å and remains stable, indicating that the global structure of the HEWL remains folded close to its native-state structure at the baseline temperature, as it should. The C_α RMSD of the 691 K replica increases over the first 20 ns as the HEWL protein unfolds, only occasionally sampling a folded state with low C_α RMSD.

Another indication of the extent of unfolding is provided by the STRuctural IDENTification (STRIDE) program²⁷ for evaluating secondary structure. Figure 9 presents the temporal evolution of the α -helix fractions in the HEWL protein for the 300 K and 691 K replicas in the TIGER2 (Figure 9(A)) and TIGER2A (Figure 9(B)) simulations. The TIGER2 result shown in Figure 9(A) is consistent with the RMSD analysis in Figure 8(A) in showing that the 300 K and 691 K replicas exhibit a similar amount of secondary structure, with the baseline replica having lost much of its α -helix content and the 691 K replica preserving too much α -helix content. The TIGER2A results are far more reasonable, with the α -helix fraction of the 300 K replica remaining stable and fluctuating around the native-state structure with about 35% α -helix content throughout the 25 ns simulation, whereas the 691 K replica undergoes unfolding, with its α -helix content decreasing to below 10%.

FIG. 7. Residue helicity at 270 K obtained from REMD (solid squares),¹ TA-MP-1 (solid circles), TA-MP-2 (open circles), T-MP-1 (solid triangles), and T-MP-2 simulations (open triangles). The residue helicities were calculated using the last 10 000 samples for each simulation.

FIG. 8. Temporal evolution of the HEWL C α RMSD (vs. crystal structure) for the 300 K and 691 K replicas obtained from TIGER2 (A) and TIGER2A (B) simulations.

FIG. 9. Temporal evolutions of HEWL α -helix fractions of the 300 K and 691 K replicas obtained from TIGER2 (A) and TIGER2A (B) simulations. The error bars represent the standard deviation for the block-averaged values over each 1.0 ns of simulation.

The sampling times were different for the TIGER2 and TIGER2A simulations for this system (10 ns vs. 20 ns, respectively). But Figures 8 and 9 indicate clearly that the improved accuracy of the TIGER2A method is not due to the greater sampling time; if the data in Figures 8(B) and 9(B) were truncated at 10 ns, the conclusions would be the same. We have included the additional data for TIGER2A, rather than obscuring it, simply because it is available. Conversely, we did not continue the TIGER2 simulations beyond 10 ns because these were the first simulations in which we noticed the problems with TIGER2 sampling; once the TIGER2 simulations were seen to give nonphysical results, they were discontinued in order to investigate the causes.

IV. CONCLUSIONS

The performance of a recently developed accelerated sampling method, TIGER2, has been examined for several systems in which statistical fluctuations in solvent energy are much larger than the solute energy differences of interest. For each of these systems, ranging from simple 1-D models up through polypeptides to proteins, and with synthetic, uncoupled solvent noise up to explicit solvent, the TIGER2 method is not able to generate Boltzmann statistics at the target baseline temperature. The sampling errors in the TIGER2 simulations are caused by the energy fluctuation of solvent in the system due to thermal motion, which causes the replica exchanges to be dominated by the large fluctuations in solvent energy, such that the algorithm does not obey the detailed balance condition.

To circumvent this problem, we developed a new algorithm named TIGER2A, which adds a sampling period at the end of the sampling cycle in which the solute atoms are fixed in place while the solvent configurations are averaged,

as a method of reducing the detrimental effects of the solvent energy fluctuations.

The TIGER2A scheme is able to substantially reduce the sampling errors from TIGER2 in each of the model systems considered, generate sampling that closely approximates Boltzmann statistics, and represent the correct physical behavior of a protein in explicitly represented aqueous solution. Furthermore, the degree of accuracy in the TIGER2A simulation can be controlled by adjusting the duration of the averaging period.

The TIGER2A method preserves the advantages of the TIGER2 method in that it uses global replica exchange to effectively overcome the replica-diffusion problem inherent in the REMD method, thus providing improved sampling efficiency. In addition, this algorithm enables the number of replicas used for a simulation to be independent of the size of the system. This provides users with the freedom to select the number of replicas that they will use in a simulation based on the computing resources at hand, thus providing an accelerated sampling method that can be applied to any sized molecular system. This feature is in distinct contrast to a REMD simulation where the minimum number of replicas required to simulate a large molecular system is determined by the system size and may exceed the computational resources that are available.

Synthetic energy noise was used in some of the model systems to simplify the simulations, isolate the cause of the sampling problems, and allow adequate statistical convergence to draw unambiguous conclusions. This approach captures the effects of the very large energy fluctuations caused by the use of explicit water, illustrating the sampling problem that TIGER2 has when used with explicit solvent without the associated computational expense. It also demonstrates that the addition of the solvent-energy averaging phase at the end

of the sampling cycle with TIGER2A substantially remedies this problem. Although this approach is unable to include correlations between solute and solvent energies, they nonetheless demonstrate the problems with TIGER2 simulations with explicit solvent and demonstrate the benefit of the new TIGER2A approach.

ACKNOWLEDGMENTS

This work was supported by “RESBIO—The National Resource for Polymeric Biomaterials” funded under NIH Grant No. P41 EB001046-01A1 and the Engineering Research Centers Program of the National Science Foundation through the Center for Advanced Engineering Fibers and Films CAEFF, funded by NSF Award No. EEC-9731680. Partial support was also provided by NIH under Grant Nos. R01 GM074511-01A1 and R01 EB006163-0131. We would also like to acknowledge helpful discussions with Dr. Nadeem Vellore, University of Utah, regarding the alanine dipeptide simulations. We thank an anonymous reviewer for pointing out the potential differences between averaged solvent energies and the solvent PMF.

¹X.-F. Li, S. J. Stuart, and R. A. Latour, *J. Chem. Phys.* **130**, 174106 (2009).

²X.-F. Li and R. A. Latour, *J. Comput. Chem.* **32**, 1091 (2011).

³Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **314**, 141 (1999).

⁴N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).

⁵A. D. MacKerell, Jr., M. Feig, and C. L. Brooks III, *J. Comput. Chem.* **25**, 1400 (2004).

⁶N. M. Henriksen, D. R. Roe, and T. E. Cheatham III, *J. Phys. Chem. B* **117**, 4014 (2013).

⁷D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik,

A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P. A. Kollman, AMBER 12, University of California, San Francisco, CA, 2012.

⁸V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, *Proteins* **65**, 712 (2006).

⁹A. E. Roitberg, A. Okur, and C. Simmerling, *J. Phys. Chem. B* **111**, 2415 (2007).

¹⁰W. Im, S. Lee Michael, and C. L. Brooks, *J. Comput. Chem.* **24**, 1691 (2003).

¹¹J. H. Chen, W. Im, and C. L. Brooks III, *J. Am. Chem. Soc.* **128**, 3728 (2006).

¹²W. L. Jorgensen, *J. Phys. Chem.* **87**, 5304 (1983).

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

¹⁴B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.* **4**, 187 (1983).

¹⁵A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, *J. Phys. Chem. B* **102**, 3586 (1998).

¹⁶L. Verlet, *Phys. Rev.* **159**, 98 (1967).

¹⁷D. Levesque and L. Verlet, *J. Stat. Phys.* **72**, 519 (1993).

¹⁸S. Nosé, *Mol. Phys.* **52**, 255 (1984).

¹⁹W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).

²⁰W. F. van Gunsteren and H. J. C. Berendsen, *J. Mol. Phys.* **34**, 1311 (1977).

²¹G. E. P. Box and M. E. Muller, *Ann. Math. Stat.* **29**, 610 (1958).

²²E. D. N. S. Abeyrathne, H. Y. Lee, and D. U. Ahn, *Poult. Sci.* **92**, 3292 (2013).

²³R. Swaminathan, V. K. Ravi, S. Kumar, M. V. S. Kumar, N. Chandra, T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).

²⁴M. Refaee, T. Tezuka, K. Akasaka, and M. Williamson, *J. Mol. Biol.* **327**, 857 (2003).

²⁵X.-F. Li, C. P. O'Brien, G. Collier, N. A. Vellore, F. Wang, R. A. Latour, D. A. Bruce, and S. J. Stuart, *J. Chem. Phys.* **127**, 164116 (2007).

²⁶W. Shalongo, L. Dugad, and E. Stellwagen, *J. Am. Chem. Soc.* **116**, 8288 (1994).

²⁷D. Frishman and P. Argos, *Proteins* **23**, 566 (1995).