THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

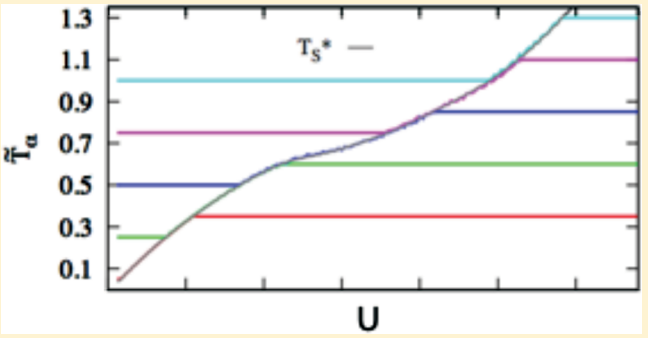
pubs.acs.org/JPCB

Replica Exchange Statistical Temperature Molecular Dynamics Algorithm

Jaegil Kim,*,† John E. Straub, and Tom Keyes*

Department of Chemistry, Boston University, Boston, Massachusetts 02215, United States

ABSTRACT: The replica exchange statistical temperature molecular dynamics (RESTMD) algorithm is presented, designed to alleviate an extensive increase of the number of replicas required as system size increases in the conventional temperature replica exchange method (*t*REM), and to obtain improved sampling in individual replicas. RESTMD optimally integrates multiple STMD (*Phys. Rev. Lett.* **2006**, *97*, 050601) runs with replica exchanges, giving rise to a flat energy sampling in each replica with a self-adjusting weight determination. The expanded flat energy dynamic sampling range allows the use of significantly fewer STMD replicas while maintaining the desired acceptance probability for replica exchanges. The computational advantages of RESTMD over conventional REM and single-replica STMD are explicitly demonstrated with an application to a coarse-grained protein model. The effect of two different kinetic temperature control schemes on the sampling efficiency is explored for diverse simulation conditions.



I. INTRODUCTION

The temperature replica exchange method (*t*REM),[1,2] or parallel tempering (PT),[3] has been gaining popularity in computer simulation of diverse complex systems with rugged energy landscapes.[4−15] Performing multiple, independent runs for a sequence of temperatures, and occasionally swapping configurations among replicas, enables a significant acceleration of configurational sampling, overcoming an ergodicity breaking problem at low temperatures.[16,17]

To maintain an appreciable acceptance probability for replica exchanges, neighboring replicas must overlap in energy. In the *t*REM, this means that the average energy separation, ΔU, between neighboring replicas should be comparable to the typical energy fluctuation, δU. Since $\Delta U = C_v \Delta T$ and $\delta U = T(C_v)^{1/2}$ in the canonical ensemble, C_v and ΔT being the heat capacity and the temperature separation of adjacent replicas, respectively, we have $\Delta U/\delta U = (\Delta T/T)(C_v)^{1/2} \sim 1$. The number of replicas is proportional to $1/\Delta T$, which is seen to increase as $\sim(N)^{1/2}$ with increasing system size.

In order to circumvent the extensive increase of replicas required for *t*REM, several sophisticated REM variants have been proposed.[18−35] One approach is to combine the generalized ensemble method (GEM)[36−38] with REM, in which each replica utilizes a non-Boltzmann sampling weight, yielding a delocalized energy distribution and allowing sufficient energy overlap with fewer replicas. The multicanonical replica exchange method (MUCAREM)[39−42] has been shown to produce comparable performance to *t*REM using half the number of replicas. However, a necessary prior weight determination is a limiting obstacle to its widespread use.

Recently, we proposed the replica exchange statistical temperature Monte Carlo (RESTMC) algorithm[43] by combining the ingredients of statistical temperature MC (STMC)[44] and REM. In RESTMC, each individual replica samples a range of temperatures with a self-adjusting weight determination and attains a flat energy distribution, leading to a significant decrease of the number of replicas with no deterioration in sampling efficiency, as demonstrated in Lennard-Jones clusters with $N = 31$, possessing a challenging double-funneled energy landscape. However, in many condensed phase simulations of complex fluids and biomolecules, molecular dynamics (MD) is preferable to MC, due to the difficulty of designing effective Monte Carlo moves in low-energy, compact states.

In this paper, we present the replica exchange statistical temperature molecular dynamics (RESTMD) algorithm, and evaluate the performance gain with respect to both conventional replica exchange MD (REMD) and single-replica STMD. In contrast to RESTMC, requiring the replica exchange of coordinates only, RESTMD must also exchange momenta. This requirement poses the challenge of selecting an optimal kinetic temperature control scheme for each replica, and here we explore two possibilities. With applications to the Honeycutt−Thirumalai coarse-grained *BLN* protein model,[45] the performance gain of RESTMD is explicitly demonstrated in terms of tunneling events, exchange acceptance rates, and inherent structures.

Special Issue: Macromolecular Systems Understood through Multiscale and Enhanced Sampling Techniques

Received: January 11, 2012
Revised: April 13, 2012
Published: April 27, 2012

   dx.doi.org/10.1021/jp300366j | *J. Phys. Chem. B* 2012, 116, 8646−8653

The paper is organized as follows: In section II, the basic formulation and the simulation protocols of RESTMD are presented. In section III, the performance of RESTMD over *t*REM or single-replica STMD is examined for the *BLN*-46-mer and 69-mer.[46] The conclusion and a brief summary are presented in section IV.

II. METHODS

A. Replica Exchange Statistical Temperature Molecular Dynamics. The ideal sampling weight in the combination of the generalized ensemble method (GEM) and REM is the reciprocal of the partial density of states, $\Omega_\alpha[U(\mathbf{x})]$, with $U(\mathbf{x})$ being the potential energy at configuration $\mathbf{x}$ and α being the replica index, as

$$W_\alpha^{\mathrm{id}}(U) = 1/\Omega_\alpha(U) \tag{1}$$

giving rise to a flat energy distribution and maximizing energy overlap between neighboring replicas. The main challenge is that Ω_α is not known *a priori*, and the corresponding estimate, $\tilde{\Omega}_\alpha$, must be determined before a production run.[42]

RESTMD is a hybrid sampling method merging multiple statistical temperature molecular dynamics (STMD) runs with replica exchanges (see refs 43 and 44 for details). Since a single STMD simulation performs a random walk in energy with a self-adjusting weight determination, seeking a flat energy distribution, joining multiple STMD runs via replica exchanges, provides a unique way to alleviate the system size dependence of conventional *t*REM while avoiding both the unknown weight dependence and the difficulty of choosing MC moves.

RESTMD replicas are characterized by the generalized ensemble weights in configurational space

$$W_\alpha = 1/\tilde{\Omega}_\alpha(U) = \exp\{-\tilde{S}_\alpha(U)\} \tag{2}$$

with $\tilde{S}_\alpha = \ln \tilde{\Omega}_\alpha$ ($k_B = 1$) being the estimate for the exact configurational entropy and $S_\alpha = \ln \Omega_\alpha$ in the microcanonical ensemble.

The key quantity in RESTMD is the replica-dependent statistical temperature, $T_\alpha(U) = [\partial S_\alpha/\partial U]^{-1}$. Instead of directly refining the extensive $\tilde{\Omega}_\alpha$ as in Wang–Landau sampling,[47] RESTMD refines the statistical temperature estimate, $\tilde{T}_\alpha(U) = [\partial \tilde{S}_\alpha/\partial U]^{-1}$, via the dynamic modification scheme

$$\tilde{T}'_{\alpha,i\pm1} = \frac{\tilde{T}_{\alpha,i\pm1}}{1 \mp \delta f \tilde{T}_{\alpha,i\pm1}} \tag{3}$$

upon a visit to discretized energy U_i, where $i = G(U/\Delta)$, Δ is the bin size, and $G(x)$ returns the nearest integer to x. The prime denotes the updated value.

As the modification factor, $\delta f = \ln f/(2\Delta)$, gradually decreases to zero via the repeated Wang–Landau operation, $f \rightarrow (f)^{1/2}$, every specified number of MD steps, $\tilde{T}_\alpha(U)$, systematically converges to the true $T_\alpha(U) = [\partial S/\partial U]^{-1}$ and a flat energy distribution is realized in each replica.

B. Molecular Dynamics Implementation. By considering the generalized ensemble in eq 2 as a canonical ensemble associated with an effective potential, $w_\alpha(U) = T_\alpha^{\mathrm{kin}}\tilde{S}_\alpha(U)$, at the fixed kinetic temperature, $T_\alpha^{\mathrm{kin}} = 1/\beta_\alpha^{\mathrm{kin}}$, the flat energy sampling weight is achieved

$$W_\alpha = \exp\{-\beta_\alpha^{\mathrm{kin}} w_\alpha(U)\} = \exp(-\tilde{S}_\alpha(U)) \tag{4}$$

RESTMD equations of motion, coupled to the Nosé–Hoover thermostat,[48] are obtained as

$$\begin{aligned}
\dot{\mathbf{q}}_i &= \mathbf{p}_i \\
\dot{\mathbf{p}}_i &= -\nabla_{\mathbf{q}_i} w_\alpha(U) - \xi \mathbf{p}_i = \gamma_\alpha(U)\mathbf{f}_i - \xi \mathbf{p}_i \\
\dot{\xi} &= [K(\mathbf{p}_i) - N_f T_\alpha^{\mathrm{kin}}]/Q
\end{aligned} \tag{5}$$

where $K(\mathbf{p}_i) = \sum_i \mathbf{p}_i^2/2$, $\gamma_\alpha(U) = T_\alpha^{\mathrm{kin}}/\tilde{T}_\alpha(U)$, and $\mathbf{q}_i$, $\mathbf{p}_i$, and $\mathbf{f}_i$ correspond to the coordinate, the momentum, and the force of the ith particle, respectively. Here, ξ and Q represent the conjugate momentum and fictional mass of the Nosé–Hoover thermostat, determining the strength of the thermal coupling to a system having N_f degrees of freedom.

Equation 5 corresponds to an ordinary molecular dynamics simulation combined with an instaneous force scaling with an energy-dependent factor, $\gamma_\alpha(U) = T_\alpha^{\mathrm{kin}}/\tilde{T}_\alpha(U)$, and the average kinetic energy is maintained at the fixed kinetic temperature T_α^{kin}. In contrast to RESTMC, requiring an explicit form of the extensive $\tilde{S}_\alpha$, eq 5 needs only the intensive $\tilde{T}_\alpha(U)$.

To restrict the dynamic sampling range of each replica, the instantaneous value of $\tilde{T}_\alpha(U)$ is always maintained between $T_\alpha^{\min}$ and $T_\alpha^{\max}$, corresponding to low and high temperature bounds, respectively, by enforcing

$$\tilde{T}_\alpha(U) = \begin{cases} T_\alpha^{\min} & \text{for } \tilde{T}_\alpha(U) \le T_\alpha^{\min} \\ T_\alpha^{\max} & \text{for } \tilde{T}_\alpha(U) \ge T_\alpha^{\max} \end{cases} \tag{6}$$

The advantage of RESTMD over conventional REMD and a single STMD is straightforward. Flat energy replicas maintain a sufficient overlap between neighbors with far fewer replicas than in conventional REMD, and the division of temperature space into smaller windows enables a significant acceleration of the weight determination process compared to single-replica STMD.

C. Kinetic Temperature Control for Replicas. While RESTMC swaps only coordinates, $\mathbf{x}$, replica exchanges in RESTMD must also consider the momenta, $\mathbf{p}$. The full sampling weight in phase space $(\mathbf{x}, \mathbf{p})$ in RESTMD is obtained as

$$W_\alpha^{\mathcal{H}} \sim \exp\{-\beta_\alpha^{\mathrm{kin}}\mathcal{H}_\alpha(\mathbf{x}, \mathbf{p})\} \tag{7}$$

where $\mathcal{H}_\alpha(\mathbf{x}, \mathbf{p})$, the Hamiltonian in the α-th replica, is equal to the sum of the kinetic energy, $K_\alpha(\mathbf{p})$, and the effective potential, $w_\alpha[U(\mathbf{x})]$.

To preserve a detailed balance, the coordinates and momenta exchange between replicas α and α', characterized by $(\mathbf{x}, \mathbf{p})$ and $(\mathbf{x}', \mathbf{p}')$, should be accepted with the probability

$$A_{\alpha\alpha'} = \min[1, \exp(\Delta_{\mathbf{x}} + \Delta_{\mathbf{p}})] \tag{8}$$

where $\Delta_x = \tilde{S}_\alpha(\mathbf{x}) + \tilde{S}_{\alpha'}(\mathbf{x}') - \tilde{S}_\alpha(\mathbf{x}') - \tilde{S}_{\alpha'}(\mathbf{x})$ and the momentum-dependent $\Delta_{\mathbf{p}}$ is

$$\beta_\alpha^{\mathrm{kin}}\{K_\alpha(\mathbf{p}) - K_\alpha(\tilde{\mathbf{p}})\} + \beta_{\alpha'}^{\mathrm{kin}}\{K_{\alpha'}(\mathbf{p}') - K_{\alpha'}(\tilde{\mathbf{p}}')\}$$

with $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{p}}'$ being the new momenta in α and α' replica after the exchange. It should be noted that $\Delta_{\mathbf{p}}$ varies with the choice of $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{p}}'$ (see below).

Depending on the temperature control scheme of the replicas, i.e., T_α^{kin}, the force scaling in eq 5 and the acceptance rule for replica exchanges in eq 8 should be handled differently. Here we explore two kinetic temperature control schemes as follows:

(i) Inhomogeneous Kinetic Temperature Control (IK): Each replica has a predetermined, different, kinetic temperature, T_α^{kin},

dx.doi.org/10.1021/jp300366j | J. Phys. Chem. B 2012, 116, 8646–8653

as usual in conventional REMD, resulting in $\gamma_\alpha(U) = T_\alpha^{kin}/\tilde{T}_\alpha(U)$. Following the original convention in replica exchange MD,[5] the exchanged momenta are scaled to assign new momenta

$$\tilde{\mathbf{p}} = \sqrt{T_\alpha/T_{\alpha'}}\,\mathbf{p}' \quad \text{and} \quad \tilde{\mathbf{p}}' = \sqrt{T_{\alpha'}/T_\alpha}\,\mathbf{p} \tag{9}$$

resulting in $\Delta_p = 0$ in eq 8.

(ii) Homogeneous Kinetic Temperature Control (HK): All replicas have the same kinetic temperature $T_\alpha^{kin} = T_0$, with $\gamma_\alpha(U) = T_0/\tilde{T}_\alpha(U)$. Since $\beta_\alpha^k = T_{\alpha'}^{kin}$, the direct exchange of momenta $\mathbf{p}$ and $\mathbf{p}'$, i.e., $\tilde{\mathbf{p}} = \mathbf{p}'$ and $\tilde{\mathbf{p}}' = \mathbf{p}$, leads to $\Delta_p = 0$, not requiring any scaling. Even though the kinetic temperature is maintained at the same T_0, each replica samples a different energy region according to eq 5. Interestingly, the acceptance probability for replica exchanges in HK also reduces to $\min[1, \exp(\Delta_x)]$.

D. Statistical-Temperature Weighted Histogram Analysis Method. After a production run, RESTMD replicas are combined to determine the inverse statistical temperature $\beta_S(U) = 1/T_S(U)$. Here we exploit the iteration-free, statistical-temperature weighted histogram analysis method (ST-WHAM),[49] a recently developed alternative to the widely used WHAM.[50]

ST-WHAM provides a unique way to determine an optimal estimate, β_S^*, for a set of simulations associated with the sampling weights, $W_\alpha(U)$, and resulting histograms, $H_\alpha(U)$, as (see ref 49 for details)

$$\begin{aligned}\beta_S^*(U) &= \sum_\alpha f_\alpha^*(U)[\beta_\alpha^H(U) + \tilde{\beta}_\alpha(U)] \\ &= \sum_\alpha f_\alpha^*(U)\beta_\alpha^*(U)\end{aligned} \tag{10}$$

where $\tilde{\beta}_\alpha(U) = 1/\tilde{T}_\alpha(U)$, $\beta_\alpha^H(U) = \partial \ln H_\alpha/\partial U$, and $f_\alpha^* = H_\alpha/\sum_\alpha H_\alpha$. Denoting $\tilde{\beta}_\alpha(U) = -\partial \ln W_\alpha/\partial U$, we identify $\beta_\alpha^* = \partial \ln(H_\alpha/W_\alpha)/\partial U = \partial \ln \Omega_\alpha^*/\partial U = \partial S_\alpha^*/\partial U$, with Ω_α^* and S_α^* being the reweighted partial density of states and entropy estimate for each replica α, respectively. Note that ST-WHAM directly determines the optimal $\beta_\alpha^*(U)$ as a weighted superposition of the replica-dependent $\beta_\alpha^*(U)$ with no iterative evaluations for partition functions intrinsic to conventional WHAM.

In an actual implementation, $\beta_\alpha^H(U)$ at U_j in eq 10 is approximated by its finite difference form, $\ln\{H_\alpha(U_{j+1})/H_\alpha(U_{j-1})\}/2\Delta$, and $\beta_S^*(U)$ is defined on discrete energy grids U_j. The corresponding entropy estimates, $S^*(U)$ and $S_\alpha^*(U)$, are obtained by substituting the corresponding $T_S^*(U)$ and $T_\alpha^*(U)$ into

$$S^*(U) = \sum_{j=L}^{i_{max}} L_j(U_{j+1}) + L_{i_{max}}(U) \tag{11}$$

where $L_j = 1/\eta_j \ln[1 + \eta_j(U - U_j)/T_j^*]$, $\eta_j = (T_{j+1}^* - T_j^*)/\Delta$. Here $i_{max} = i - 1$ for $U \in [U_i - \Delta/2, U_i]$ and $i_{max} = i$ for $U \in [U_i, U_i + \Delta/2]$.

Once $S^*(U)$ is determined, any canonical thermodynamic property at an arbitrary temperature β is obtained by reweighting as

$$A(\beta) = \sum_U A(U)P_\beta(U) \tag{12}$$

where $P_\beta(U) = e^{-\beta(U-S^*)}/\sum_U e^{-\beta(U-S^*)}$.

E. Simulation Protocols. In our studies so far, STMD uses the velocity-Verlet integrator and the Nosé–Hoover thermostat, although other choices could be implemented equally well. Since $\tilde{T}_\alpha(U)$ is being dynamically updated in the weight refining stage of RESTMD, trajectories are initially not in full equilibrium. Data are not taken until δf approaches zero, and detailed balance is recovered with a correct estimation of $\tilde{T}_\alpha(U)$. Practical simulation protocols of RESTMD are outlined as follows:

(i) Divide the temperature range of interest, between T_{min} and T_{max}, into M overlapping windows, each covering the range $[T_\alpha^{min}, T_\alpha^{max}]$. The overlap between neighboring windows is adjusted by the overlap parameter κ as $T_\alpha^{min} = T_\alpha - \kappa(T_{\alpha+1} - T_\alpha)$ and $T_\alpha^{max} = T_\alpha + \kappa(T_{\alpha+1} - T_\alpha)$. A set of discrete temperatures, T_α ($\alpha = 1, ..., M + 1$), is sequentially distributed as $T_{min}(T_{max}/T_{min})^{(\alpha-1)/(M-1)}$ in a geometric allocation, and $T_{min} + [(\alpha - 1)/(M - 1)](T_{max} - T_{min})$ in an equidistant allocation. Select the simulation parameters, Δ and f, and the kinetic temperature control scheme. Initialize $\tilde{T}_\alpha(U) = (T_\alpha^{min} + T_\alpha^{max})/2$ and set $T_\alpha^{kin} = T_\alpha$ for the case of inhomogeneous temperature control.

(ii) Run preliminary STMD simulations to obtain a rough estimate for $\tilde{T}_\alpha(U)$ with no replica exchanges and with the modification factor fixed at the initial value, until the minimum and maximum values of the estimate, $\tilde{T}_\alpha(U)$, reach T_α^{min} and T_α^{max}, respectively.

(iii) Propagate multiple STMD runs with replica exchange attempts at specified intervals, using the acceptance rule of eq 8. Depending on the kinetic temperature control scheme, the assignment of new momenta after the replica exchange should be done differently as described in section 2C. Reduce the modification factor as $f \to (f)^{1/2}$ every fixed $(M \times N_r)$ MD steps. In contrast to the single STMD, where the reduction of f occurs only when the energy histogram satisfies a strict flatness condition,[44] RESTMD exploits a periodic reduction scheme, since the dynamic sampling range of each replica is significantly reduced and replica exchanges assist the system in avoiding trapping and reaching ergodicity.

(iv) Move to a production run with the frozen $\tilde{T}_\alpha(U)$ once $\delta f \leq 10^{-8}$. Simulation results from all replicas are joined to estimate $T_S^*(U)$ and $S^*(U)$ via ST-WHAM.[49]

III. SIMULATION RESULTS AND DISCUSSIONS

We have chosen the 46-residue, off-lattice Honeycutt–Thirumalai *BLN* protein model,[45] denoted *BLN*-46mer, as a benchmark for RESTMD. This model has been studied extensively,[51−58] with methods including single-replica STMD,[44] and provides a good example of a rugged energy landscape. We used the same potential form and parameter set as a reference.[55]

We first performed several RESTMD simulations subject to the inhomogeneous kinetic (IK) temperature control scheme with varying numbers of replicas, M = 5, 10, 20, and 30. Selecting $T_{min} = 0.05$ and $T_{max} = 1.3$, the temperature range spanned by each replica was determined via the equidistant allocation scheme, with a varying overlap parameter, κ. The energy bin size is chosen as $\Delta = 1$. The modification factor f was reduced to $(f)^{1/2}$, starting from 1.0001, every $N_r = (2.5 \times 10^6)/M$ MD steps in each replica. After 10 reductions of f, reaching $\delta f \approx 5 \times 10^{-8}$, the $\tilde{T}_\alpha$ are effectively converged, and

dx.doi.org/10.1021/jp300366j | J. Phys. Chem. B 2012, 116, 8646–8653

the production data were collected for 2×10^8 additional MD steps. Note that the weight determination time is effectively fixed at 2.5×10^7 MD steps regardless of M.

For comparison, we also performed conventional REMD simulation using 30 replicas for the same temperature range. Due to the vanishing acceptance of replica exchanges at low temperatures, the geometric allocation scheme was applied to generate T_α, and replica temperatures are assigned as $(T_\alpha + T_{\alpha+1})/2$ (α = 1, ..., M), corresponding to the kinetic temperatures in RESTMD conjugated with the inhomogeneous temperature control (RESTMD-IK). The global minimum configuration was used as an initial configuration for all replicas, and replica exchanges were attempted every 10^2 MD steps in each replica in both RESTMD and REMD simulations.

The converged statistical temperature estimates, $\tilde{T}_\alpha(U)$, with M = 5 and κ = 0.2 in Figure 1a, illustrate the characteristic

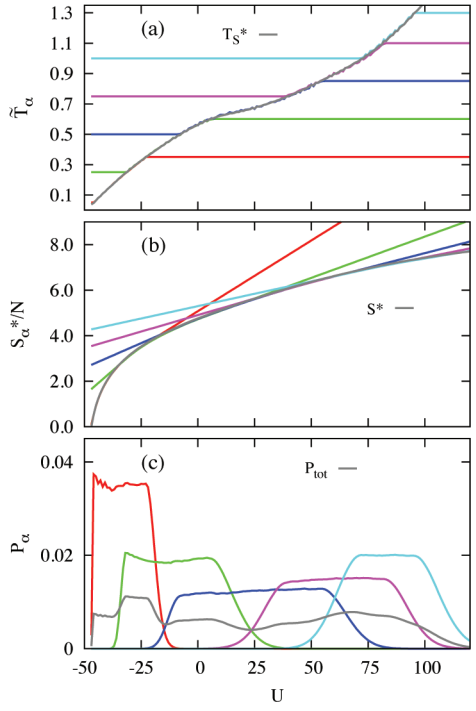


Figure 1. *BLN*-46mer: (a) $\tilde{T}_\alpha(U)$ and $T_S^*(U)$, (b) $S_\alpha^*(U)$ and $S^*(U)$, and (c) $P_\alpha(U)$ and $P_{tot}(U)$ for the RESTMD-IK run with M = 5 and κ = 0.2. Bottom to top, α = 1, 2, 3, 4, and 5 at U = 100 in (a) and the same color scheme is applied for all figures. Both T_S^* and S^* are determined by ST-WHAM for the production run for 2×10^8 MD steps.

features of RESTMD. Replica-dependent $\tilde{T}_\alpha(U)$ associated with different temperature windows smoothly join together across the overlapping regions. The superimposed statistical temperatures shows a monotonous increase with increasing U and a characteristic slope variation, indicating the collapse transition region ($T_\theta \approx 0.65$), and are in good agreement with $T_S^*(U)$ determined by ST-WHAM.

The replica-dependent, reweighted entropy estimates, $S_\alpha^*(U)$, in Figure 1b, determined from $\tilde{T}_\alpha$ and H_α via eq 11, almost coincide with the optimal S^* for the nonvanishing $f_\alpha^*(U)$. The flat energy distribution found in $P_\alpha(U)$ for $T_\alpha^{min} \le \tilde{T}_\alpha(U) \le T_\alpha^{max}$ in Figure 1c indicates that the difference between the estimated $\tilde{T}_\alpha(U)$ and $T_S(U)$ is negligible. In the low and high energy region of each replica, RESMTD samples the canonical ensembles at T_α^{min} and T_α^{max}, respectively, giving rise to the Gaussian decay in $P_\alpha(U)$ and the linear $S_\alpha^*(U)$ for $\tilde{T}_\alpha(U) < T_\alpha^{min}$ and $\tilde{T}_\alpha(U) > T_\alpha^{max}$.

Due to the restricted sampling range in each replica, the weight determination process in RESTMD is significantly shortened in comparison to the single STMD, in which the modification factor f decreases only when the histogram satisfies a flatness criterion for the whole energy region. Additionally, of course, the stochastic process of replica exchange assists the system in avoiding trapping and reaching ergodicity. As illustrated in Figure 2a, $\tilde{T}_\alpha(U)$ associated with f =

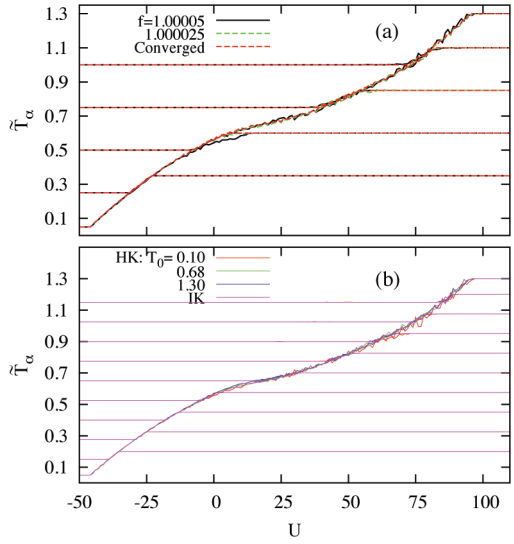


Figure 2. *BLN*-46mer: (a) $\tilde{T}_\alpha(U)$ at different f values in the RESTMD-IK run with M = 5 and κ = 0.2 and (b) converged $\tilde{T}_\alpha(U)$ in RESTMD-HK with T_0 = 0.1, 0.68, and 1.3 and RESTMD-IK for M = 10 and κ = 0.2.

1.00005 at 1.25×10^6 MD steps shows a minimal deviation at mid temperatures, and already becomes indistinguishable from the converged function at f = 1.000025. In comparison to the single STMD ($\approx 5 \times 10^7$ MD steps),[44] the weight determination time has been shortened to one-half ($\approx 2.5 \times 10^7$ MD steps) in RESTMD.

We also performed RESTMD simulations using the homogeneous temperature control (RESTMD-HK), with different kinetic temperatures, T_0 = 0.1, 0.68, and 1.3, for M = 10 and κ = 0.2. The resulting $\tilde{T}_\alpha(U)$ are (Figure 2b) in good agreement with each other (except for T = 0.1, which exhibits a small deviation with a rugged behavior in the high energy region), and collapse on that of RESTD-IK, implying that our algorithm provides the same thermodynamic properties, mostly encoded by $\tilde{T}_\alpha$, irrespective of the temperature control scheme (see also Figure 7b).

The advantage of the RESTMD algorithm over REMD is explicitly demonstrated in energy and replica trajectories of two arbitrary chosen replicas in Figure 3a and b. Even with only five replicas (M = 5), RESTMD shows very frequent round-trips in both energy and replica space. With increasing M, Figure 3c and d, tunneling (round-trip) transitions become less frequent. This slowing down is mainly attributed to the increase of required replica exchanges for a tunneling event. On the other hand, conventional REMD shows a transient trapping at low energy regions even with the geometric temperature allocation, and round trips are much less frequent. The advantage of RESTMD over REMD is particularly pronounced at low

dx.doi.org/10.1021/jp300366j | J. Phys. Chem. B 2012, 116, 8646−8653

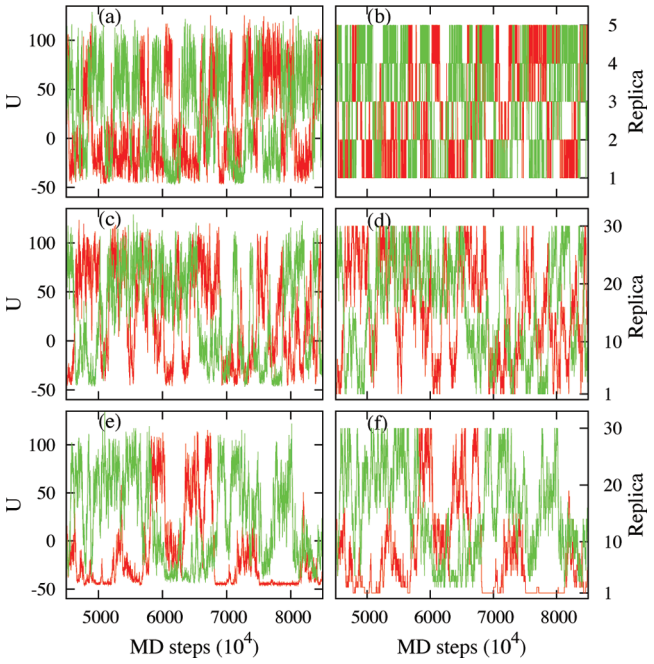


Figure 3. *BLN*-46mer: (a) Energy and (b) replica trajectories of two representative replicas in RESTMD-IK with M = 5 and κ = 0.2, (c) energy and (d) replica trajectories in RESTMD-IK with M = 30 and κ = 0.2, and (e) energy and (f) replica trajectories in REMD with M = 30.

temperatures, where REMD requires a dense distribution of replicas with closely spaced temperatures to counteract the vanishing acceptance of exchanges.

The success of RESTMD stems from its capacity to retain (or enhance) the acceptance of replica exchanges with fewer (or the same) number of replicas. The average acceptance probabilities, $\bar{A}_\alpha = \langle A_{\alpha,\alpha+1}\rangle$, with $\langle ...\rangle_\alpha$ being the ensemble average, are plotted as a function of α in RESTMD-IK runs with varying κ from −0.2 to 0.4 at fixed M = 10 in Figure 4a.

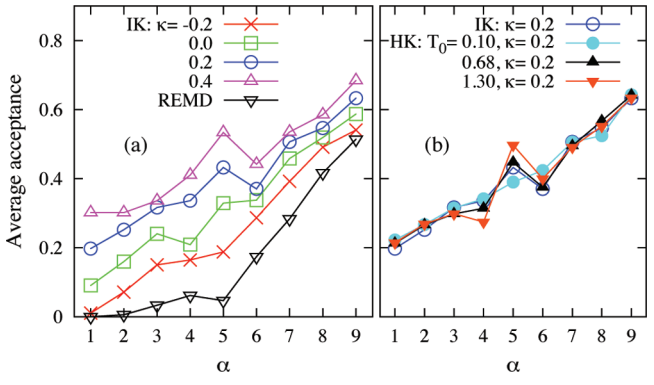


Figure 4. *BLN*-46mer: Average acceptance probabilities in (a) RESTMD-IK simulations with varying κ at fixed M = 10 and REMD with M = 10 and (b) RESTM-HK simulations with M = 10 and κ = 0.2.

The enlarged energy overlap with increasing κ enables a systematic enhancement of $\bar{A}_\alpha$ especially at low temperatures, while $\bar{A}_\alpha$ in conventional REMD with the same M rapidly diminishes at low temperatures.

We also compared $\bar{A}_\alpha$ of RESTMD with homogeneous temperature control (RESTMD-HK) and various T_0 in Figure 4b. Except for some intermediate replicas, α = 4−5 near the transition region $T_\theta \approx 0.65$, the average acceptances collapse on those of RESTMD-IK. Since the average acceptance is directly determined by an overlap integral between adjacent replicas,[34] a slight elevation of $\bar{A}_S$ with increasing T_0 implies that the sampling dynamics around the transition region could be affected by the choice of T_0. Consistently, RESTMD-HK with $T_0 \approx T_\theta$ shows a profile of $\bar{A}_\alpha$ most similar to that of RESTMD-IK across the replicas.

To examine the performance gain more quantitatively, we compared the accumulated tunneling transitions,[14,36,59] N_U, measured between U = −42 and −95, in Figure 5a. In

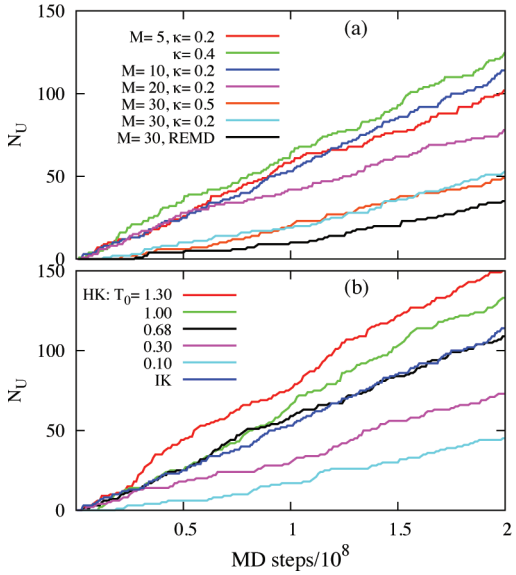


Figure 5. *BLN*-46mer: Accumulated tunneling transitions, N_U, in energy space (a) for RESTMD-IK runs with varying M and κ and REMD with M = 30 and (b) for RESTMD-HK runs at different T_0 for M = 10 and κ = 0.2. For comparison, N_U of RESTMD-IK is also plotted in (b).

RESTMD-IK simulations with varying M and fixed κ = 0.2, fewer replicas, maintaining sufficient overlaps with flat energy distributions, enable far more frequent tunneling transitions than conventional REMD with M = 30. A similar trend in N_U is also observed with increasing κ for fixed M, but its effect is less dramatic. The best result in RESTMD-IK runs shows about a 3-fold enhancement of N_U over conventional REMD.

The interesting observation is that the rate of tunneling transitions monotonically increases in proportion to the reference kinetic temperature, T_0, in RESTMD-HK. As demonstrated in Figure 5b, N_U shows a significant impairment at the low T_0 = 0.1, but it begins to rise with increasing T_0, attaining its highest value at T_0 = 1.3. RESTMD-HK with T_0 = 0.68 ≈ T_θ exhibits the N_U profile most similar to that of RESTM-IK, just as with $\bar{A}_\alpha$ in Figure 4b. The systematic enhancement of N_U with increasing T_0 arises from enhanced diffusion in energy space with an elevated kinetic temperature.

To investigate the effect of T_0 on N_U in RESTMD-HK more clearly, we compute a bias in acceptance probabilities, defined as $B_\alpha^\pm = N_\alpha^\pm/(N_\alpha^+ + N_\alpha^-)$, with N_α^+ and N_α^- being the number of transitions from α to $(\alpha + 1)$ and from α to $(\alpha - 1)$, respectively. $B_\alpha^\pm$ measures a bias of each walker in replica space and reduces to 0.5 for a perfect random walker.

dx.doi.org/10.1021/jp300366j | J. Phys. Chem. B 2012, 116, 8646−8653

As demonstrated in Figure 6a and b, $B_\alpha^\pm$ in the RESTMD-HK with T_0 = 0.1 is almost 0.5 across all replicas except for both

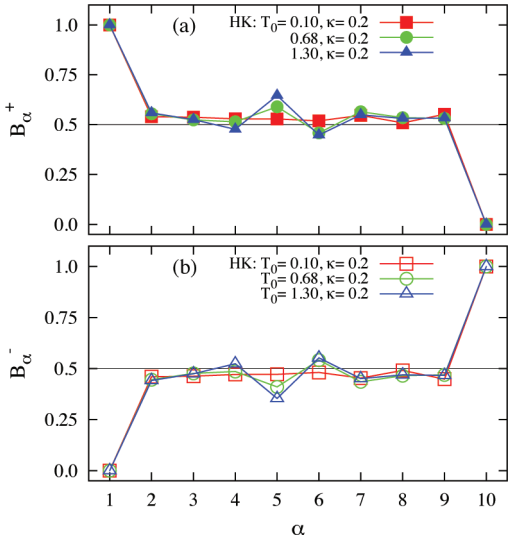


Figure 6. *BLN*-46mer: Bias in acceptance probabilities: (a) B_α^+ and (b) B_α^- for RESTMD-HK runs with varying T_0 and the RESTMD-IK run for M = 10 and κ = 0.2. The horizontal line corresponding to a perfect random walk in both (a) and (b) is plotted for visualization.

ends, implying that the dynamics in replicas is almost random. On the other hand, a significant bias is observed for the RESTMD-HK with T_0 = 1.3 at α = 5 and 6. Both the increase of B_5^+ and decrease of B_6^- from 0.5 reveal that the dynamics of the fifth and sixth replicas is tuned to facilitate tunneling transitions with increasing T_0. Indeed, the temperature ranges of these two replicas are directly associated with the protein collapse region around $T_\theta \approx 0.65$ as $(T_5^{min}, T_5^{max}) = (0.525, 0.7)$ and $(T_6^{min}, T_6^{max}) = (0.65, 0.825)$, implying that the enhanced configurational sampling by the elevated kinetic temperature is the main booster for the enhanced tunneling transitions.

For diverse simulation conditions such as varying M and κ, heat capacities, C_v, of all RESTMD-IK runs for 2×10^8 MD steps collapse on the same curve in Figure 7a, and are in good agreement with those of REMD. Except for T_0 = 0.1, associated with the poorest N_U, the same thermodynamic consistency is also observed between RESTMD-IK and RESTMD-HK with varying T_0 in Figure 7b.

In addition to the applications to equilibrium sampling, RESTMD is also useful in finding a global minimum in complex energy landscapes. To explore this possibility, we applied REMSTD to the *BLN*-69mer[46] characterized by the deeply, multifunneled energy landscape,[44,58] arising from the conformal diversity generated by hydrophobic mismatches between six β strands. We performed several RESTMD-IK runs with varying M and κ = 0.5 and conventional REMD simulation between T_{min} = 0.05 and T_{max} = 1.3. Instead of the ground state configuration with U = −99.189, an arbitrary chosen extended configuration was used for the initial configuration in all replicas. The initial modification factor was set to f = 1.0005 to speed the search for low energy states, and the explored energy landscape is characterized by inherent structures (IS)[60] determined by a conjugate-gradient minimization algorithm.

The IS plot of the lowest replica as a function of the total simulation time in Figure 8 clearly illustrates a superior power of RESTMD in finding the global minimum. For a smaller

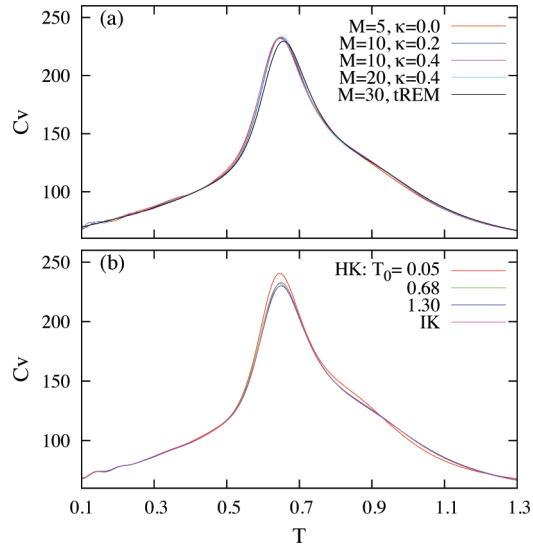


Figure 7. *BLN*-46mer: Heat capacities C_v determined by (a) RESTMD-IK runs with varying M and κ and (b) RESTMD-HK runs at different T_0 values.

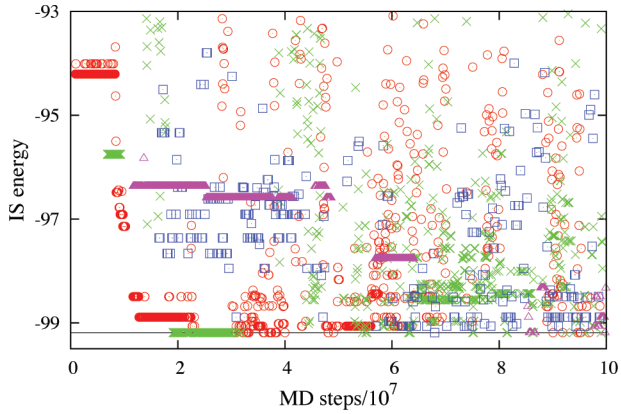


Figure 8. *BLN*-69mer: Inherent structure energies of the lowest replica in RESTMD and REMD simulations. Circles, RESTMD with M = 10 and κ = 0.2; crosses, RESTMD with M = 10 and κ = 0.5; squares, RESTMD with M = 30 and κ = 0.5; triangles, REMD with M = 30. The horizontal line near the bottom indicates the ground state energy, −99.189.

number of replicas, M = 10, RESTMD finds the known global minimum at ∼1.91×10^7 and ∼3.02×10^7 MD steps for κ = 0.5 and 0.2, respectively, and RESTMD with M = 30 and κ = 0.5 reaches the ground state at ∼6.81×10^7 MD steps, while conventional REMD takes ∼8.57×10^7 MD steps. Fewer steps are required in all cases if all replicas are considered, but RESTMD retains its advantage. In addition to the accelerated search for the global minimum, RESTMD shows broadly scattered IS energies in Figure 8, implying that several low-lying IS states are sampled even in the single replica, and transitions among those IS states are very frequent through replica exchanges, while the IS energy profiles of REMD are narrow and transitions among IS states are very rare.

IV. CONCLUSION

The replica exchange statistical temperature molecular dynamics (RESTMD) algorithm has been developed to mitigate the extensive increase of the number of replicas with increasing system size in conventional REMD. RESTMD

dx.doi.org/10.1021/jp300366j | J. Phys. Chem. B 2012, 116, 8646−8653

combines several STMD runs with replica exchanges, in which each replica samples a range of temperatures and achieves a flat energy sampling with a self-adjusting sampling weight characteristic to the replica-dependent statistical temperature. In contrast to conventional REMD, the systematic enhancement of energy overlaps between neighboring RESTMD replicas allows a significant reduction of the number of replicas while maintaing effective configurational sampling.

The quantitative performance comparison between RESTMD, conventional REMD, and single-replica STMD for the coarse-grained protein model, *BLN*-46mer, subject to a significant degree of energetic frustration, reveals that RESTMD provides a considerable enhancement in the rate of convergence of simulations accompanied with accelerated tunneling transitions. It is also shown that the narrowed temperature window of each replica enables a significant reduction of the weight determination time in RESTMD with the periodic reduction scheme of *f*, contrary to the histogram-flatness reduction scheme in single STMD.

We also explored two different kinetic temperature control schemes in RESTMD associated with a different assignment rule for the momenta after the replica exchange. It is found that RESTMD-HK yields more frequent tunneling transitions than RESTMD-IK with increasing reference kinetic temperature, but both implementations provide the same thermodynamic properties at a moderate reference kinetic temperature. The superior performance of RESTMD over REMD in finding a global minimum is also demonstrated in the *BLN*-69mer possessing a deep, multifunneled energy landscape.

The extensive increase of the number of replicas and a deterioration of sampling efficiency with increasing system size in conventional REMD is a significant challenge for the conformational sampling of many complex systems. This is particularly pressing in simulations of biomolecules in explicit water or in lipid bilayers, where a vast dynamic energy range, mostly contributed from solvent−solvent interactions, significantly hampers the weight determination in the generalized ensemble method. We hope that the accelerated, self-adjusting weight determination in the RESTMD algorithm, combined with the periodic reduction of *f*, will contribute to overcoming this difficulty.

An essential step toward our broad goal of exploiting STMD in realistic applications is merging it with a biosimulation package. We have created STMD-CHARMM and are currently developing STMD-NAMD. We are also testing versions which seek a flat energy distribution in the biomolecular energy only, mitigating the problem of solvent−solvent energy dominance. These single-replica applications, running on machines anywhere on a network or in the cloud, will be organized into an extremely powerful RE simulation with a master script, targeting the most challenging biophysical problems.

AUTHOR INFORMATION

Corresponding Author

*E-mail: jaegilkim89@gmail.com (J.K.); keyes@bu.edu (T.K.).

Present Address

[†]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the NSF (CHE-0750309, CHE-1114676, and CHE-0833605) and the NIH (R01 GM076688) for support.

REFERENCES

(1) Swendsen, R. H.; Wang, J. S. *Phys. Rev. Lett.* **1986**, *57*, 2607.
(2) Geyer, C. J.; Thompson, A. *J. Am. Stat. Assoc.* **1995**, *90*, 909.
(3) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604.
(4) Hansmann, U. H. *Chem. Phys. Lett.* **1997**, *281*, 140.
(5) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141.
(6) Zhou, R.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12777.
(7) Garcia, A, E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13898.
(8) Paschek, D.; Gnanakaran, S.; Garcia, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6765.
(9) Yamamoto, R.; Kob, W. *Phys. Rev. E* **2000**, *61*, 5473.
(10) Flenner, E.; Szamel, G. *Phys. Rev. E* **2006**, *73*, 061505.
(11) Widom, M.; Ganesh, P.; Kazimirov, S.; Louca, D.; Mihalkovic, M. *J. Phys.: Condens. Matter* **2008**, *20*, 114114.
(12) Liu, H.; Jordan, K. D. *J. Phys. Chem. A* **2005**, *109*, 5203.
(13) Frantsuzov, P. A.; Mandelshtam, V. A. *Phys. Rev. E* **2005**, *72*, 037102. Mandelshtam, V. A.; Frantsuzov, P. A. *J. Chem. Phys.* **2006**, *124*, 204511.
(14) Poulain, P.; Calvo, F.; Antoine, R.; Broyer, M.; Dugourd, P. *Phys. Rev. E* **2006**, *73*, 056704.
(15) Bellesia, G.; Shea, J. E. *J. Chem. Phys.* **2009**, *130*, 145103; *131*, 111102.
(16) Newman, M. E.; Barkema, G. T. *Monte Carlo Methods in Statistical Physics*; Clarendon: Oxford, U.K., 1999.
(17) Earl, D. J.; Deem, M. W. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3910.
(18) Fukunishi, H.; Watanabe, O.; Takada, S. *J. Chem. Phys.* **2002**, *116*, 9058.
(19) Whitefield, T. W.; Bu, L.; Straub, J. E. *Physica A* **2002**, *305*, 157.
(20) Jang, S.; Shin, S.; Park, Y. *Phys. Rev. Lett.* **2003**, *91*, 058305.
(21) Kim, J.; Fukunishi, Y.; Nakamura, H. *J. Chem. Phys.* **2004**, *121*, 1626. Kim, J.; Fukunishi, Y.; Kidera, A.; Nakamura, H. *J. Chem. Phys.* **2004**, *121*, 5590.
(22) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13749. Liu, P.; Huang, X.; Zhou, R.; Berne, B. J. *J. Phys. Chem. B* **2006**, *110*, 19018.
(23) Cheng, X.; Cui, G.; Hornak, V.; Simmering, C. *J. Phys. Chem. B* **2005**, *109*, 8220.
(24) Lyman, E.; Ytreberg, F. M.; Zuckerman, M. D. *Phys. Rev. Lett.* **2006**, *96*, 028105.
(25) Trebst, S.; Troyer, M.; Hansmann, U. H. *J. Chem. Phys.* **2006**, *124*, 174903.
(26) Calvo, F. *J. Chem. Phys.* **2006**, *123*, 124106.
(27) Rick, S. W. *J. Chem. Phys.* **2007**, *126*, 054102.
(28) Liu, P.; Voth, G. A. *J. Chem. Phys.* **2007**, *126*, 045106.
(29) Kamberaj, H.; van der Vaart, A. *J. Chem. Phys.* **2007**, *127*, 234102.
(30) Brenner, P.; Sweet, C. R.; Von Handorf, D.; Izaguirre, J. A. *J. Chem. Phys.* **2007**, *126*, 074103.
(31) Zhang, C.; Ma, J. *Phys. Rev. E* **2007**, *76*, 036708.
(32) Ballard, A. J.; Jarzynski, C. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 12224.
(33) Kar, P.; Nadler, W.; Hansmann, U. H. *Phys. Rev. E* **2009**, *80*, 056703.
(34) Kim, J.; Straub, J. E. *J. Chem. Phys.* **2009**, *130*, 144114.
(35) Kim, J.; Keyes, T.; Straub, J. E. *J. Chem. Phys.* **2010**, *132*, 224107.
(36) Berg, B. A.; Celik, T. *Phys. Rev. Lett.* **1991**, *69*, 2292; *Phys. Lett. B* **1991**, *267*, 249.
(37) Hansmann, U. H.; Okamoto, Y. *Phys. Rev. E* **1997**, *56*, 2228.
(38) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymer* **2001**, *60*, 96.
(39) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **2000**, *329*, 261.
(40) Calvo, F.; Doye, J. P. *Phys. Rev. E* **2000**, *63*, 010902.
(41) Faller, R.; Yan, Q.; de Pablo, J. J. *J. Chem. Phys.* **2002**, *116*, 5419.

(42) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *J. Chem. Phys.* **2003**, *118*, 6664, 6676.
(43) Kim, J.; Keyes, T.; Straub, J. E. *J. Chem. Phys.* **2009**, *130*, 124112.
(44) Kim, J.; Straub, J. E.; Keyes, T. *Phys. Rev. Lett.* **2006**, *97*, 050601; *J. Chem. Phys.* **2007**, *126*, 135101; *Phys. Rev. E* **2007**, *76*, 011913.
(45) Honeycutt, J. D.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3526. Guo, Z.; Thirumalai, D.; Honeycutt, J. D. *J. Chem. Phys.* **1992**, *97*, 525.
(46) Larrass, S. A.; Pegram, L. M.; Gordon, H. L.; Rothstein, S. M. *J. Chem. Phys.* **2003**, *119*, 13149.
(47) Wang, F.; Landau, D. P. *Phys. Rev. Lett.* **2001**, *86*, 2050; *Phys. Rev. E* **2001**, *64*, 056101.
(48) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.
(49) Kim, J.; Keyes, T.; Straub, J. E. *J. Chem. Phys.* **2011**, *135*, 061103.
(50) Ferrenberg, A. M.; R. H. Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195.
(51) Guo, Z.; Thirumalai, D. *Biopolymer* **1985**, *36*, 83.
(52) Guo, Z.; Brooks, C. L., III. *Biopolymer* **1987**, *42*, 745.
(53) Miller, M. A.; Wales, D. J. *J. Chem. Phys.* **1999**, *111*, 6610.
(54) Komatsuzaki, T.; Hoshino, K.; Matsunaga, Y.; Rylance, G. J.; Johnston, R. L.; Wales, D. J. *J. Chem. Phys.* **2005**, *122*, 084714.
(55) Lee, Y. H.; Berne, B. J. *J. Phys. Chem. A* **2000**, *104*, 86.
(56) Kim, J.; Keyes, T. *J. Phys. Chem. B* **2008**, *112*, 954.
(57) Kim, S. Y. *J. Chem. Phys.* **2010**, *133*, 135102.
(58) Oakley, M. T.; Wales, D. J.; Johnston, R. L. *J. Phys. Chem. B* **2011**, *115*, 11525.
(59) Dayal, P.; Trebst, S.; Wessel, S.; Wurtz, D.; Troyer, M.; Sabhapandit, S.; Coppersmith, S. N. *Phys. Rev. Lett.* **2004**, *92*, 097201.
(60) Stillinger, F. H.; Weber, T. A. *Phys. Rev. A* **1983**, *28*, 2408; *Science* **1984**, *225*, 983.