Accelerating atomic-level protein simulations by flat-histogram techniques

Sigurður Æ. Jónsson, Sandipan Mohanty, and Anders Irbäck

Citation: J. Chem. Phys. **135**, 125102 (2011); doi: 10.1063/1.3643328
View online: http://dx.doi.org/10.1063/1.3643328
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v135/i12
Published by the American Institute of Physics.

Accelerating atomic-level protein simulations by flat-histogram techniques

Sigurður Æ. Jónsson,[1,a)] Sandipan Mohanty,[2,b)] and Anders Irbäck[1,c)]
[1]*Computational Biology and Biological Physics, Lund University, Sölvegatan 14A, SE-223 62 Lund, Sweden*
[2]*Institute for Advanced Simulation, Jülich Supercomputing Centre, Forschungszentrum Jülich, D-52425 Jülich, Germany*

(Received 29 July 2011; accepted 7 September 2011; published online 29 September 2011)

Flat-histogram techniques provide a powerful approach to the simulation of first-order-like phase transitions and are potentially very useful for protein studies. Here, we test this approach by implicit solvent all-atom Monte Carlo (MC) simulations of peptide aggregation, for a 7-residue fragment (GIIFNEQ) of the Cu/Zn superoxide dismutase 1 protein (SOD1). In simulations with 8 chains, we observe two distinct aggregated/non-aggregated phases. At the midpoint temperature, these phases coexist, separated by a free-energy barrier of height 2.7 k_BT. We show that this system can be successfully studied by carefully implemented flat-histogram techniques. The frequency of barrier crossing, which is low in conventional canonical simulations, can be increased by turning to a two-step procedure based on the Wang-Landau and multicanonical algorithms. © *2011 American Institute of Physics*. [doi:10.1063/1.3643328]

I. INTRODUCTION

Generalized-ensemble techniques have become a widely recognized tool for speeding up statistical-mechanical simulations of systems with complex free-energy landscapes. For protein folding and aggregation studies, the currently most widely used among these techniques is replica exchange,[1–3] also called parallel tempering, which requires very little preprocessing and is ideally suited for parallel computation. A closely related method is simulated tempering.[4,5]

The replica exchange and simulated tempering methods have been successfully used to study small polypeptide chains, which typically do not show a pronounced two-state folding behavior. However, for systems exhibiting a first-order-like phase transition, these methods are expected to fail. To overcome this problem, generalizations of the methods have been proposed.[6–9]

Another possibility is to turn to flat-histogram methods such as the multicanonical algorithm.[10] The first step in a multicanonical calculation is to estimate the density of states, $g(E)$, where E denotes internal energy. Having obtained this estimate, $\tilde{g}(E)$, one simulates the ensemble defined by the microstate probability distribution $P_\nu \propto 1/\tilde{g}(E_\nu)$, in which the distribution of E is approximately uniform. Finally, properties of the original system are recovered by means of reweighting techniques.[11–14]

The first multicanonical study of the folding of a small peptide was reported already several years ago,[15] and previous applications of this method also include coarse-grained simulations of peptide aggregation.[16] However, the algorithm has not gained the same popularity as the convenient replica exchange method, in part because of the required estimation of $g(E)$.

An important step forward was the development by Wang and Landau[17,18] of a simple and general scheme for this task. The Wang-Landau method has been applied to a variety of problems, including the phase structure of long homopolymer chains.[19,20]

It has been shown that flat-histogram techniques such as the multicanonical and Wang-Landau methods can be useful for atomic-level protein simulations as well.[15,21–25] However, none of the protein systems studied so far displayed a first-order-like phase transition. Therefore, the full potential of this approach remains incompletely explored.

In this article, we study the aggregation of a 7-residue peptide by implicit solvent all-atom MC simulations with 8 chains. We use this system as a testbed for a simulation procedure based on flat-histogram techniques. For comparison, we also carry out canonical constant-temperature simulations of the same system.

The peptides are found to form β-sheet-containing aggregates at low temperatures, while being disordered and non-aggregated at high temperatures. At the midpoint temperature, the two phases coexist, as manifested by a bimodal energy distribution.

For complex systems such as proteins, the bottom part of the energy landscape may consist of narrow minima that are difficult to sample and not necessarily low in *free* energy, at the temperatures of interest. In this situation, flat-histogram techniques must be implemented with care, because a uniform sampling in energy all the way down to the lowest lying level might be both costly and unnecessary. Otherwise, one risks slowing down the simulations through a time-consuming exploration of narrow minima with a negligible occupancy at biologically relevant temperatures.

Our simulations of the above mentioned peptide system focus on the transition between the aggregated and non-aggregated phases. To study this transition, we introduce a generalized ensemble, in which the energy distribution is flat in the coexistence range. Outside this range, the energy

a)Electronic mail: sigurdur.aegir@thep.lu.se.
b)Electronic mail: s.mohanty@fz-juelich.de.
c)Electronic mail: anders@thep.lu.se.

distribution falls off rapidly, to avoid unnecessary sampling of low energies. Our calculations consist of two steps. The first step serves to determine this generalized ensemble, by means of the Wang-Landau method. In the second step, corresponding to a multicanonical production run, the generalized ensemble is simulated using standard MC techniques.

II. METHODS

A. Simulated ensemble

Our simulation procedure assumes that the system of interest displays two distinct phases as a function of temperature, and that the energy distribution is bimodal at the midpoint temperature, which we will denote by T_{m}. Let E_1 and E_2 denote the energies at which the two peaks are centered.

The procedure amounts to first constructing and then simulating a generalized ensemble, the definition of which is illustrated in Fig. 1. The microstate probability distribution is given by $P_\nu \propto 1/\gamma(E_\nu)$, where the function $\gamma(E)$ (see Fig. 1(a)) is defined by

$$\gamma(E) \propto \begin{cases} g(E) & \text{if } E_1 \le E \le E_2, \\ \exp(E/k_{\mathrm{B}}T_{\mathrm{m}}) & \text{otherwise,} \end{cases} \tag{1}$$

where k_{B} is Boltzmann's constant. With this P_ν, the probability distribution of E (see Fig. 1(b)) becomes

$$P(E) \propto \begin{cases} 1 & \text{if } E_1 \le E \le E_2, \\ g(E)\exp(-E/k_{\mathrm{B}}T_{\mathrm{m}}) & \text{otherwise.} \end{cases} \tag{2}$$

For brevity, this ensemble will be referred to as the "$1/\gamma$" ensemble. It is defined so as to have a flat $P(E)$ between E_1 and E_2, to promote transitions between the two phases. Outside this range, $P(E)$ is proportional to the canonical energy distribution at $T = T_{\mathrm{m}}$. This choice ensures that low energies will not be more extensively sampled than what is needed to characterize the system at $T = T_{\mathrm{m}}$.

To determine $\gamma(E)$, one has to estimate $g(E)$ for $E_1 \le E \le E_2$. To this end, we use the Wang-Landau method.[17,18] In principle, this calculation can be restricted to the interval $E_1 \le E \le E_2$. In practice, a slightly larger interval must be used, because E_1 and E_2 are *a priori* unknown.

Having determined $\gamma(E)$, we simulate the $1/\gamma$ ensemble by standard MC methods. From this simulation, canonical averages are extracted by use of reweighting techniques.[12]

In the $1/\gamma$ ensemble, $P(E)$ is thus flat in the coexistence region, while $P(E) \propto P_{\mathrm{can}}(E)$ outside this range. This may be compared with the recently proposed well-tempered ensemble,[26] where the energy distribution is $P_{\mathrm{can}}(E)^{1/a}$ for some $1 < a < \infty$, corresponding to a global but finite flattening of the canonical energy distribution.

B. The Wang-Landau algorithm

This section provides a brief outline of the Wang-Landau method,[17,18] followed by a description of two non-standard choices we made when implementing it.

The Wang-Landau algorithm successively builds up an estimate, $\tilde{g}(E)$, of the density of states, $g(E)$. Usually, the function $\tilde{g}(E)$ is initially set to 1 for all energies (or energy bins). After each step in the simulation, $\tilde{g}(E)$ is increased by a factor $f > 1$, $\tilde{g}(E) \to f\tilde{g}(E)$, for the current energy E, while it is left unchanged for all other E. The simulation is controlled by an accept/reject question, where the acceptance probability is given by

$$P_{\mathrm{acc}}(\nu \to \nu') = \min\left[1, \frac{\tilde{g}(E_\nu)}{\tilde{g}(E_{\nu'})}\right]. \tag{3}$$

Since P_{acc} depends on the current $\tilde{g}(E)$, detailed balance is fulfilled only in the limit $f \to 1$.

The modification factor f is typically assigned an initial value of $f = \mathrm{e}$. During the course of the simulation, f is gradually decreased toward 1, by changing $f \to \sqrt{f}$ whenever a certain criterion is met. Usually, this criterion has the form[17,18]

$$\min_E h(E) > \alpha\,\overline{h(E)}, \tag{4}$$

where $h(E)$ is a histogram of the energies visited since the last change of f, $\alpha < 1$ is a parameter, and $\overline{h(E)}$ denotes an instantaneous average of $h(E)$ over E. The simulation is continued until $\ln f < \epsilon$, where $\epsilon > 0$ is a stopping parameter.

The convergence properties of this scheme were analyzed by several groups. One important result is that the error in the estimated density of states at a given f, scales as $\sqrt{\ln f}$.[27] Furthermore, it was demonstrated that f should be decreased as $1/t$ with simulation time t, rather than exponentially, in order to avoid asymptotic saturation of the error for small f.[28] Methods addressing the error saturation problem have been devised.[28–31] In the present study, we stick to the fast exponential update $f \to \sqrt{f}$, because we use the Wang-Landau

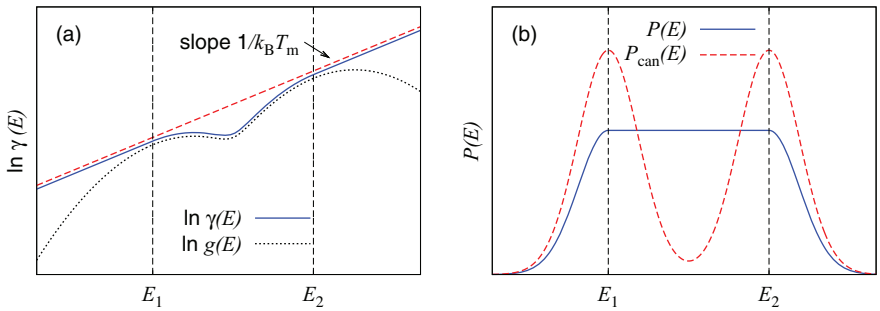

FIG. 1. Schematic illustration of the simulated ensemble. (a) The definition of the function $\gamma(E)$ (see Eq. (1)), which determines the microstate probability distribution, $P_\nu \propto 1/\gamma(E_\nu)$. (b) The probability distribution of energy, $P(E)$ (see Eq. (2)), along with the canonical energy distribution at $T = T_{\mathrm{m}}$, $P_{\mathrm{can}}(E)$.

method to prepare for multicanonical production runs, rather than to generate final results.

Our implementation of the Wang-Landau method is essentially as described above, but with two modifications. The first modification is in the criterion for when to change f. Instead of using Eq. (4), we update f when the number of tunneling events since the last change of f, n_t, satisfies

$$n_t > \tau \tag{5}$$

where τ is a threshold parameter. A tunneling event is a traversal of the whole energy spectrum, from one end to the other. The tunneling frequency is a key characteristic of the dynamics.[32] We were led to replace Eq. (4) by Eq. (5) because of results from toy model calculations, as described in the Appendix.

Our second modification lies in the initial values of $\tilde{g}(E)$ and of the parameter f. At the early stages of preliminary runs started with $f = \mathrm{e}$, virtually no β-structure was seen. Instead, the system was driven into kinetically more easily accessible α-helical low-energy states. A balanced sampling of α- and β-structure was observed only after f had been reduced several times. In the calculations below, we therefore start with $\ln f = 2^{-10}$, instead of $\ln f = 1$. Because of a relatively small step-size parameter $\ln f$, we further use an improved initial guess for $\ln \tilde{g}(E)$, $\ln \tilde{g}(E) = E/k_B T_m +$ constant, in place of $\ln \tilde{g}(E) =$ constant. Having a rough estimate of the *a priori* unknown parameter T_m is sufficient for this linear *ansatz* to be useful.

C. Peptide model

The peptide we study is a 7-residue fragment, GIINFEQ, of the protein SOD1. This fragment is part of the second β-strand in native SOD1 (residues 16–22; Protein Data Bank code 1AZV). A link between SOD1 and amyotrophic lateral sclerosis (ALS) exists, as over 100 ALS-associated mutations in SOD1 have been identified.[33]

We simulate a system of 8 GIINFEQ peptides enclosed in a periodic box of size $(126\ \text{Å})^3$. All simulations are started from random initial conditions, with different random number seeds in different runs.

Our simulations are based on an implicit solvent all-atom model with torsional degrees of freedom, which has been described in detail in Refs. 34 and 35. In short, the potential is composed of four terms, $E = E_{loc} + E_{ev} + E_{hb} + E_{sc}$. The first term, E_{loc}, contains local interactions between atoms separated by only a few covalent bonds. The other three terms are non-local in character: E_{ev} represents excluded-volume effects, E_{hb} is a hydrogen-bond potential, and E_{sc} describes residue-specific interactions, based on hydrophobicity and charge, between pairs of side chains. Energies quoted below are given in a unit corresponding to $\sim$1.33 kcal/mol.

This potential was developed through folding thermodynamics studies of a structurally diverse set of peptides and small proteins, while deliberately keeping it as simple as possible.[35] It is worth noting that the same set of parameters is used for α, β, as well as α/β proteins. Previous applications of this model include aggregation studies of the 42-residue amyloid β-peptide[36] and of several short peptides.[37,38]

D. MC details

We simulate GIINFEQ aggregation using MC dynamics. Five different elementary moves are employed: (i) rotations of individual backbone angles, (ii) a semi-local backbone update, biased Gaussian steps, which rotates eight consecutive angles simultaneously,[39] (iii) rotations of individual side-chain angles, (iv) rigid-body translations of whole chains, and (v) rigid-body rotations of whole chains. The relative frequencies of the updates (i)–(v) are 25 %, 12 %, 47 %, 8 %, and 8 %, respectively.

All our simulations are carried out using the open source package PROFASI,[40] to which a new routine for Wang-Landau simulations was added. Statistical errors quoted below are calculated using the jackknife method.[41]

III. RESULTS AND DISCUSSION

The simulated system of 8 GIIFNEQ peptides in a periodic box displays two distinct phases: a disordered and non-aggregated high-temperature phase, and a low-temperature phase in which the peptides form β-sheet-containing aggregated structures. The two phases coexist in a narrow temperature range, where the energy distribution is bimodal (see below). Examples of aggregated structures can be seen in Fig. 2. A vast majority of the observed low-energy conformations share a common overall β-sandwich topology. However, there is no single dominant aggregated structure, because the β-strand organization varies, as illustrated by Fig. 2.

In preliminary runs with 2 and 4 GIIFNEQ peptides, at identical peptide concentration, the same phase structure was observed, but the free-energy barrier between the two coexisting phases was lower for these system sizes. In what follows, we focus on the 8-chain system, as a challenging testbed for the two-step simulation procedure described in Sec. II.

The first step of the procedure is to determine the function $\gamma(E)$ of Eq. (1), and thus construct the $1/\gamma$ ensemble illustrated in Fig. 1. To this end, we estimate the density of states $g(E)$ for $25 \le E \le 165$ through a set of 16 independent Wang-Landau runs. The runs are started with a modification factor of $\ln f = 2^{-10}$ and stopped when $\ln f < 2^{-23}$. Figure 3(a) shows the MC time evolution of the energy E in a typical Wang-Landau run. The tunneling frequency is high in the beginning of the run, where the system tends to be driven away from the last visited region. This bias decreases as $\ln f$

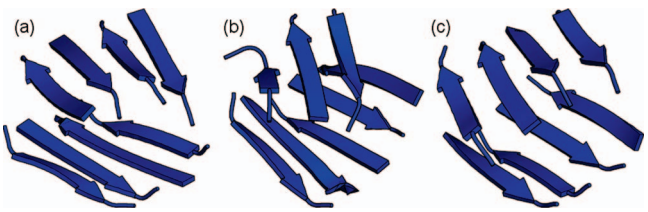

FIG. 2. Snapshots from the simulations showing typical low-energy conformations for the system of 8 GIIFNEQ peptides. The structures share an overall β-sandwich topology, but differ in the organization of the strands. The energies are (a) $E \approx -21.6$, (b) $E \approx -21.2$, and (c) $E \approx -20.2$.

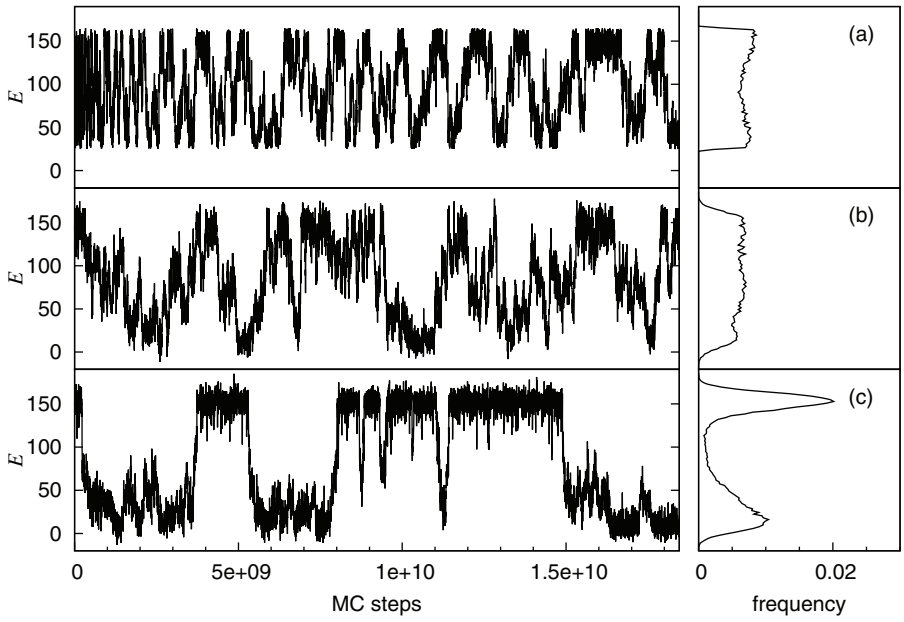

FIG. 3. MC time evolution of the energy in typical runs with three different methods for the system of 8 GFIINEQ peptides. The right panel shows histograms of energies visited in the respective runs. Each run required ~470 core hours on a 2.26 GHz Nehalem processor. For each of the three methods, 16 independent runs of this length were generated. (a) Wang-Landau simulation, (b) simulation of the $1/\gamma$ ensemble (see Fig. 1), and (c) canonical-ensemble simulation at $k_B T = 0.496$.

is reduced. As a result, the tunneling frequency goes down, but tunneling events still continue to occur throughout the run.

The coexistence range $E_1 \leq E \leq E_2$ can be identified through an indentation in the calculated $\ln g(E)$, as indicated in Fig. 1. This way we estimate $E_1 \approx 28$, $E_2 \approx 155$, and $k_B T_m \approx 0.4960$. These parameters along with the calculated shape of $g(E)$ for $E_1 \leq E \leq E_2$ determine the function $\gamma(E)$.

Knowing $\gamma(E)$, we next simulate the $1/\gamma$ ensemble by standard MC techniques. To collect statistics, a set of 16 independent runs is generated in this case as well. A representative run-time trajectory can be seen in Fig. 3(b). The distribution of energies visited in the run is approximately flat in the interval $E_1 \leq E \leq E_2$, as it should be. Figure 4 displays the canonical energy distribution, $P(E)$, at $k_B T = 0.496$, as obtained from the $1/\gamma$-ensemble simulations by reweighting.[12] The bimodality of $P(E)$ shows that the transition between the two phases indeed is first-order-like. A small asymmetry in peak height indicates that the true T_m is slightly higher than what we estimated above based on the Wang-Landau runs ($k_B T_m \approx 0.4960$). The improved estimate provided by the $1/\gamma$-ensemble simulations is $k_B T_m \approx 0.4967$. At $T = T_m$, we find that $P(E)$ is suppressed by a factor 15 in the valley between the two equally high peaks, which corresponds to a free-energy barrier of height $2.7\, k_B T$.

To validate and assess the efficiency of this two-step simulation procedure, we now compare the results above with data from a set of 16 conventional canonical-ensemble simulations at $k_B T = 0.496$. The $P(E)$ distribution extracted above turns out to be in perfect agreement with the results obtained from these control simulations, as can be seen from Fig. 4, thus confirming the validity of the two-step procedure. Compared to the canonical-ensemble simulations, it can further be seen that the results from the $1/\gamma$-ensemble simulations have smaller statistical errors, despite that the runs are equally long in both cases. The use of the $1/\gamma$ ensemble thus yields a more efficient sampling.

Another even more direct way to see this is to study the run-to-run variation of some observable. Figure 5 shows normalized single-run estimates of the heat capacity, C_v, at $k_B T = 0.496$, for both sets of simulations. In both cases, each data point represents one of 16 independent runs. The spread of the data is large for the canonical-ensemble simulations. It

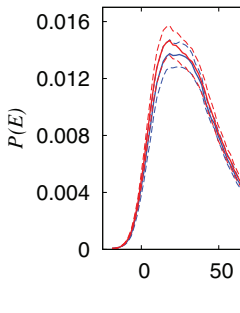

FIG. 4. Probability distribution of E for the system of 8 GFIINEQ peptides at $k_B T = 0.496$, as obtained from simulations of the $1/\gamma$ (blue) and canonical (red) ensembles, respectively. Dashed lines indicate statistical 1σ errors.

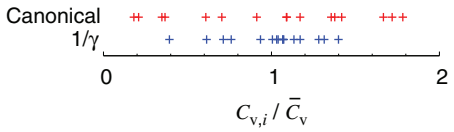

FIG. 5. Run-to-run variation of the heat capacity $C_v = (\langle E^2 \rangle - \langle E \rangle^2)/k_B T^2$, calculated at $k_B T = 0.496$, in simulations of the canonical (red) and $1/\gamma$ (blue) ensembles. The single-run estimates, $C_{v,i}$, are normalized by the mean, $\bar{C}_v$. The length of the runs is the same with both methods.

is smaller for the $1/\gamma$-ensemble simulations, which confirms that sampling is more efficient in this case.

The origin of this speedup is evident from the run-time trajectories in Fig. 3. In the canonical run (Fig. 3(c)), the system is stuck in the same state for long periods, which in particular makes the relative population of the two phases statistically difficult to estimate. This problem can, at least in part, be overcome by the use of the $1/\gamma$ ensemble (Fig. 3(b)). Compared to the canonical runs, we find that the tunneling frequency, on average, is a factor 2.8 higher in the $1/\gamma$-ensemble simulations, where intermediate energies are not statistically suppressed. The tunneling frequency is a key figure when studying surface effects[16,42] and properties that strongly depend on the relative population of the coexisting states, such as the heat capacity.

A critical property of the above two-step procedure is that unnecessary sampling of low energies is avoided. In the production runs, low energies are suppressed by the weight factor $1/\gamma(E)$. In the preparatory Wang-Landau runs, it is sufficient to cover the coexistence range $E_1 \leq E \leq E_2$. In our Wang-Landau runs, the sampled energy range was $25 \leq E \leq 165$, whereas the ground-state energy is < -20 for this system (see Fig. 2). In preliminary calculations, we observed a notably ($\gtrsim$50 %) slower convergence upon a moderate extension of the energy range to $5 \leq E \leq 165$. A further extension down to the ground-state level would have made the calculations much more time-consuming.

Finally, let us mention that we also performed replica exchange simulations of the same system, using a set of 16 temperatures between 0.42 and 0.56 (distributed so as to have a constant overlap between neighboring energy distributions). Here, the tunneling frequency was lower by factors 4.0 and 11.2 compared to our canonical and $1/\gamma$ simulations, respectively. In peptide folding studies, Okamoto and co-workers found the multicanonical method to be more efficient than replica exchange, which in turn was more efficient than canonical-ensemble methods.[43–45] When applied to our very different system, the order of the methods need not be same. On the other hand, it is likely that the poor tunneling frequency observed in our replica exchange simulations could have been improved by increasing the minimum temperature (0.42) to a value closer to T_m ($\approx$0.4967). Even more interesting, however, would be to try instead one of the replica exchange variants specifically meant for first-order-like phase transition,[6–8] but that is beyond the scope of the present article.

IV. SUMMARY AND OUTLOOK

We have implemented and tested a two-step procedure for protein simulations, based on flat-histogram techniques. Our test system exhibits a first-order-like transition between two aggregated/non-aggregated phases. As far as we know, this is the first time flat-histogram techniques have been used for atomic-level simulations of a protein system at phase coexistence.

In the proposed approach, we construct and simulate an ensemble, $P_v \propto 1/\gamma(E_v)$, whose energy distribution is flat only in the coexistence range $E_1 \leq E \leq E_2$, while falling off rapidly outside this range. For convenience, we used a single parameter, T_m, to set the shape of the tails of the distribution. The exact shape of the tails is unimportant. The main point is to avoid unnecessary sampling of low energies.

In our simulations of the $1/\gamma$ ensemble, we find that the tunneling frequency is improved by a factor 2.8, compared to canonical-ensemble simulations. This is the speed-up factor we expect for long accurate simulations, when the relative cost of the preparatory Wang-Landau runs becomes negligible. This factor is not very large, but can still be helpful if the simulations require weeks or months to complete. Moreover, it is important to remember that the free-energy barrier of the present system, although clear, is only moderately high (2.7 $k_B T$). Experimentally estimated folding/unfolding barriers are often higher than this. With a higher barrier, we expect the speed-up factor to increase.

Because of the suppression of low energies, our method alone is not well suited for ground-state searches, but it may be used to generate a diverse set of starting points for fast local minimizations by, for instance, the conjugate gradient method. It should be pointed out, however, that knowledge of the ground state of a protein model need not be thermodynamically important, because the ground state may be insignificantly populated throughout the limited temperature range which is biologically relevant.

A property unanticipated by us was the artificial preference for α-structure over β-structure seen for large values of the Wang-Landau parameter f. The formation of β-structure, which requires the establishment of long-range contacts, seemed to be hampered by the bias away from the last visited region present at large f. This problem might be general and show up in simulations of other protein systems as well. Due to this problem, we found it advantageous to use a relatively small initial f.

The aggregation of the peptide studied here, GIIFNEQ, has, to our knowledge, not been investigated experimentally. However, the aggregation prediction program WALTZ (Ref. 46) points at this part of the SOD1 protein as particularly prone to form amyloid structure.[47] Our observation of β-sheet-rich aggregated structures is consistent with this prediction.

Let us finally stress that the flat-histogram calculations presented here were carried out without exhausting the toolbox of possible refinements. One possibility we did not

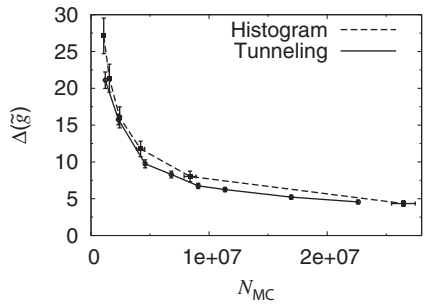

FIG. 6. Average error in the estimated density of states against average number of steps required for convergence, in toy model Wang-Landau simulations with the histogram-based (Eq. (4)) and tunneling (Eq. (5)) criteria. Different data points correspond to different α and τ, respectively.

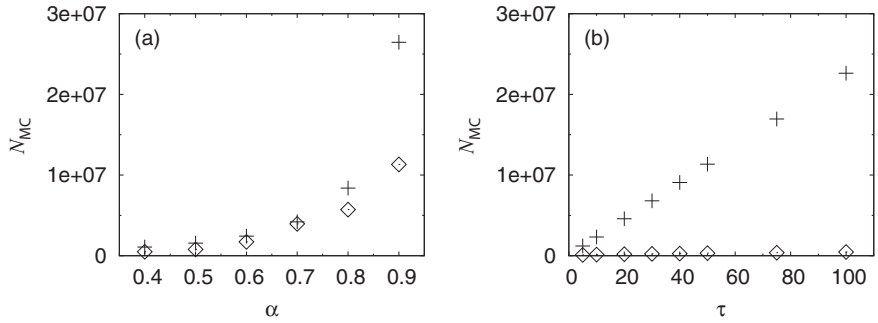

FIG. 7. Mean (+) and standard deviation (⋄) of the number of steps required for convergence, N_{MC}, in toy model simulations with two different criteria for when to change the Wang-Landau parameter f. (a) The histogram criterion, Eq. (4). (b) The tunneling criterion, Eq. (5).

explore is the use of control variables other than the energy. Another possible improvement would be to fine-tune the shape of the energy distribution in the coexistence range so as to maximize the tunneling frequency, rather than prescribing a flat distribution.[48,49]

ACKNOWLEDGMENTS

The simulations were performed at the LUNARC facility, Lund University. This work was in part funded by the Swedish Research Council.

APPENDIX: TOY MODEL ANALYSIS OF THE *f* PARAMETER

We mentioned two different criteria, given by Eqs. (4) and (5), for when to change the Wang-Landau modification factor f. In this appendix, we compare these criteria by simulations of a simple toy model.

We consider a system with $N+1$ possible states ν. Each state ($\nu = 0, \ldots, N$) is assigned an energy

$$E_\nu = \frac{E_{\max}\nu}{N} \tag{A1}$$

and an entropy

$$S_\nu = \begin{cases} S_{\max}\nu/N - 2S_b\nu/N & \text{if } \nu \le N/2, \\ S_{\max}\nu/N - 2S_b(1-\nu/N) & \text{if } \nu > N/2, \end{cases} \tag{A2}$$

where $E_{\max}$, $S_{\max}$, and S_b are parameters (all assumed positive). In the calculations below, we set $N = 99$, $E_{\max} = 25$ (a.u.), $S_{\max}/k_B = 25$, and $S_b/k_B = 12$.

At $T = T_m = E_{\max}/S_{\max}$, the free energy $F_\nu = E_\nu - TS_\nu$ is tent shaped, with two degenerate minima at $\nu = 0$ and $\nu = N$. The minima are separated by a barrier of height $T_m S_b$ centered at $\nu = N/2$.

An elementary Wang-Landau update of this system can be defined as follows. If the system is in state ν, a new state $\nu' = \nu$ or $\nu \pm 1$ is proposed with probability $\min(1, e^{(S_{\nu'} - S_\nu)/k_B})/2$ for $\nu' = \nu \pm 1$. This proposal is accepted or rejected, with a probability of acceptance given by Eq. (3).

Using this dynamics, we examine the convergence times with the criteria, Eqs. (4) and (5), respectively, through simulations for a broad range of values of the parameters α and τ. For each parameter value, a set of $\ge$128 independent runs is generated. Each run is started with $\ln f = 1$ and stopped when $\ln f < 10^{-7}$. Two quantities are recorded: the number of elementary updates required for convergence, N_{MC}, and the deviation $\Delta(\tilde{g})$ of the final estimate $\tilde{g}(E)$ from the true density of states (e^{S_ν/k_B}), calculated as

$$\Delta(\tilde{g}) = \sum_\nu \left| \ln \frac{\tilde{g}(E_\nu)}{g(E_\nu)} \right|. \tag{A3}$$

Figure 6 summarizes the results of these simulations, for different α and τ, by showing how the error $\Delta(\tilde{g})$ varies with N_{MC}. From this figure, it can be seen that the average computational effort required to achieve a given accuracy is slightly lower with the tunneling criterion, Eq. (5), than it is with the histogram-based criterion, Eq. (4), but the difference is small.

Figure 7 shows the mean and standard deviation of N_{MC} as functions of α and τ. The standard deviation of N_{MC} is strikingly smaller with Eq. (5) than it is with Eq. (4). Although the average computational effort is similar with both criteria, we thus find that the run-to-run variation is much smaller with the tunneling criterion. Because of this robustness, we decided to use the tunneling criterion, Eq. (5), in our peptide simulations. The toy model analysis was repeated in the absence of a free-energy barrier ($S_b = 0$), with similar results.

[1] R. H. Swendsen and J. S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).
[2] C. J. Geyer and E. A. Thompson, J. Am. Stat. Assoc. **90**, 909 (1995).
[3] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
[4] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[5] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992).
[6] T. Neuhaus and J. S. Hager, Phys. Rev. E **74**, 036702 (2006).
[7] T. Neuhaus, M. P. Magiera, and U. H. E. Hansmann, Phys. Rev. E **76**, 045701 (2007).
[8] J. Kim and J. E. Straub, J. Chem. Phys. **133**, 154101 (2010).
[9] J. Kim, T. Keyes, and J. E. Straub, J. Chem. Phys. **132**, 224107 (2010).
[10] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).
[11] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).
[12] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).
[13] J. Ferkinghoff-Borg, Eur. Phys. J. B **29**, 481 (2002).
[14] M. K. Fenwick, J. Chem. Phys. **129**, 125106 (2008).
[15] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **14**, 1333 (1993).
[16] C. Junghans, M. Bachmann, and W. Janke, J. Chem. Phys. **128**, 085103 (2008).
[17] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[18] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).
[19] M. P. Taylor, W. Paul, and K. Binder, J. Chem. Phys. **131**, 114907 (2009).

[20]D. T. Seaton, T. Wüst, and D. P. Landau, Phys. Rev. E **81**, 011802 (2010).
[21]N. Rathore, T. A. Knotts IV, and J. J. de Pablo, J. Chem. Phys. **118**, 4285 (2003).
[22]C. Junghans and U. H. E. Hansmann, Int. J. Mod. Phys. C **17**, 817 (2006).
[23]T. Nagasima, A. R. Kinjo, T. Mitsui, and K. Nishikawa, Phys. Rev. E **75**, 066706 (2007).
[24]C. Gervais, T. Wüst, D. P. Landau, and Y. Xu, J. Chem. Phys. **130**, 215106 (2009).
[25]T. Yoda, Y. Sugita, and Y. Okamoto, Biophys. J. **99**, 1637 (2010).
[26]M. Bonomi and M. Parrinello, Phys. Rev. Lett. **104**, 190601 (2010).
[27]C. Zhou and R. N. Bhatt, Phys. Rev. E **72**, 025701 (2005).
[28]R. E. Belardinelli and V. D. Pereyra, Phys. Rev. E **75**, 046701 (2007).
[29]R. E. Belardinelli and V. D. Pereyra, J. Chem. Phys. **127**, 184105 (2007).
[30]C. Zhou and J. Su, Phys. Rev. E **78**, 046705 (2008).
[31]A. D. Swetnam and M. P. Allen, J. Comput. Chem. **32**, 816 (2011).
[32]P. Dayal, S. Trebst, S. Wessel, D. Würtz, M. Troyer, S. Sabhapandit, and S. N. Coppersmith, Phys. Rev. Lett. **92**, 097201 (2004).
[33]J. Valentine, P. Doucette, and S. Potter, Annu. Rev. Biochem. **74**, 563 (2005).
[34]A. Irbäck, B. Samuelsson, F. Sjunnesson, and S. Wallin, Biophys. J. **85**, 1466 (2003).
[35]A. Irbäck, S. Mitternacht, and S. Mohanty, PMC Biophys. **2**, 2 (2009).
[36]S. Mitternacht, I. Staneva, T. Härd, and A. Irbäck, J. Mol. Biol. **410**, 357 (2011).
[37]A. Irbäck and S. Mitternacht, Proteins **71**, 207 (2008).
[38]D. Li, S. Mohanty, A. Irbäck, and S. Huo, PLoS Comput. Biol. **4**, e1000238 (2008).
[39]G. Favrin, A. Irbäck, and F. Sjunnesson, J. Chem. Phys. **114**, 8154 (2001).
[40]A. Irbäck and S. Mohanty, J. Comput. Chem. **27**, 1548 (2006).
[41]R. G. Miller, Biometrika **61**, 1 (1974).
[42]C. Junghans, M. Bachmann, and W. Janke, Europhys. Lett. **87**, 40002 (2009).
[43]A. Mitsutake, Y. Sugita, and Y. Okamoto, J. Chem. Phys. **118**, 6664 (2003).
[44]A. Mitsutake, Y. Sugita, and Y. Okamoto, J. Chem. Phys. **118**, 6676 (2003).
[45]Y. Sugita and Y. Okamoto, Biophys. J. **88**, 3180 (2005).
[46]S. Maurer-Stroh, M. Debulpaep, N. Kuemmerer, M. Lopez de la Paz, I. C. Martins, J. Reumers, K. L. Morris, A. Copland, L. Serpell, L. Serrano, J. W. H. Schymkowitz, and F. Rousseau, Nat. Methods **7**, 237 (2010).
[47]M. Oliveberg, Nat. Methods **7**, 187 (2010).
[48]S. Trebst, D. A. Huse, and M. Troyer, Phys. Rev. E **70**, 046701 (2004).
[49]F. A. Escobedo and F. J. Martinez-Veracoechea, J. Chem. Phys. **129**, 154107 (2008).