Statistical-Temperature Monte Carlo and Molecular Dynamics Algorithms

Jaegil Kim,[*] John E. Straub, and Thomas Keyes

Department of Chemistry, Boston University, Boston, Massachusetts 02215, USA

A simulation method is presented that achieves a flat energy distribution by updating the statistical temperature instead of the density of states in Wang-Landau sampling. A novel molecular dynamics algorithm (STMD) applicable to complex systems and a Monte Carlo algorithm are developed from this point of view. Accelerated convergence for large energy bins, essential for large systems, is demonstrated in tests on the Ising model, the Lennard-Jones fluid, and bead models of proteins. STMD shows a superior ability to find local minima in proteins and new global minima are found for the 55 bead *AB* model in two and three dimensions. Calculations of the occupation probabilities of individual protein inherent structures provide new insights into folding and misfolding.

PACS numbers: 05.10.−a, 02.70.Rr, 87.18.Bb

The Wang-Landau (WL) Monte Carlo (MC) algorithm [1] has sparked considerable interest [1–3] due to its ability to generate a flat energy distribution. A distinguishing feature is that the density of states $\Omega(E)$ is dynamically constructed during the simulation. On visiting states with energy E, the running estimate $\tilde{\Omega}(E)$ is multiplied by the factor $f > 1$, which forces the system to visit less explored energy regions through the bias in the acceptance probability of $\min\{1, \tilde{\Omega}[E(\mathbf{r})]/\tilde{\Omega}[E(\mathbf{r}')]\}$ for a move with $\mathbf{r} \rightarrow \mathbf{r}'$, and enables a fast convergence compared to other flat energy histogram methods [4].

Two obstacles to further exploitation of the WL approach are that the discrete representation of $\Omega(E)$ on an energy grid causes the number of bins to increase extensively with system size, and the absence of a molecular dynamics (MD) version limits its applicability to complex systems where effective MC moves are not available, e.g., biomolecules. In this Letter we propose two algorithms that surmount these limitations by exploiting the correspondence between the density of states and the statistical temperature $T(E)$. The idea is to achieve a flat energy distribution by the systematic refinement of the statistical-temperature estimate $\tilde{T}(E)$ rather than the density of states estimate $\tilde{\Omega}(E)$. A robust dynamic modification scheme is proposed to transform an initially constant $\tilde{T}(E)$ to the true statistical temperature $T(E)$ via a nonuniform scaling operation that is intrinsically nonlocal and concurrently refines $\tilde{\Omega}(E)$ at the visited state and in its neighborhood. Updating the intensive variable $\tilde{T}(E)$ allows a continuum description for the entropy estimate regardless of the energy bin size, and yields substantial acceleration of the convergence with larger, and hence fewer bins, while maintaining statistical accuracy.

Sampling with the statistical weight, $e^{-\tilde{S}(E)} = 1/\tilde{\Omega}(E)$, $\tilde{S}(E)$ being the entropy estimate, can be obtained [5] from MD trajectories generated by the effective potential $T_0\tilde{S}(E)$ at the fixed temperature T_0. The dynamic modification of $\tilde{T}(E)$ and force scaling in the MD algorithm yields a energy distribution that converges to a flat histogram and effi-

ciently realizes a random walk in energy space via systematic repeated heating and cooling [6]. The performance of the statistical-temperature MC (STMC) and MD (STMD) algorithms is validated in simulations of the Ising model, the Lennard-Jones fluid, and coarse-grained protein models.

Our approach is based on the thermodynamic relation between the microcanonical entropy, $S(E) = \ln\Omega(E)$ ($k_B = 1$), and the inverse temperature, $\beta(E) = 1/T(E)$,

$$S(E) = \int^E \beta(E')dE'. \tag{1}$$

Because the entropy is uniquely determined up to a constant as a functional of $T(E)$ in Eq. (1), it is natural to seek a sampling scheme driven by the dynamic modification of the statistical temperature rather than the density of states or entropy. We introduce the running estimate for the temperature as

$$\tilde{\beta}(E) = 1/\tilde{T}(E) = \partial\tilde{S}/\partial E, \tag{2}$$

where $\tilde{S}(E) = \ln\tilde{\Omega}(E)$. The transformation of $\tilde{T}(E)$ to $T(E)$ is derived by applying the WL algorithm to the finite difference form of Eq. (2). On an equally spaced energy grid $E_j = G(E/\Delta)\Delta$, with bin size Δ and $G(x)$ returning the nearest integer to x, the WL operation of $\tilde{\Omega}_j \rightarrow f\tilde{\Omega}_j$ reduces to $\tilde{S}_j \rightarrow \tilde{S}_j + \ln f$ for a visit to energy E_j. We combine this operation with the finite difference approximation, $\partial\tilde{S}/\partial E|_{E=E_j} = \tilde{\beta}_j = 1/\tilde{T}_j \simeq (\tilde{S}_{j+1} - \tilde{S}_{j-1})/2\Delta$, and obtain the dynamic update of the inverse temperature as $\tilde{\beta}'_{j\pm 1} = \tilde{\beta}_{j\pm 1} \mp \delta f$, where $\delta f = \ln f/2\Delta$ and the prime denotes the updated value. Then

$$\tilde{T}'_{j\pm 1} = \alpha_{j\pm 1}\tilde{T}_{j\pm 1}, \tag{3}$$

where $\alpha_{j\pm 1} = 1/(1 \mp \delta f\tilde{T}_{j\pm 1})$. Some properties of Eq. (3) are that the scaling operations of decreasing $\tilde{T}_{j-1}$ and increasing $\tilde{T}_{j+1}$ transform $\tilde{T}(E)$ so that it converges to the monotonically increasing function $T(E)$; the factor $\alpha_{j\pm 1}$ approaches unity at low temperature, allowing a

fine-tuning of $\tilde{T}(E)$; the "edge effect" [7] can be avoided
by restricting updates to $T_l < \tilde{T}_j < T_h$, and maintaining
$\tilde{T}_j = T_l (T_h)$ beyond the lower (upper) temperature bounds
$T_l (T_h)$.

An STMC simulation requires the computation of $\tilde{S}(E)$
from $\tilde{T}(E)$ for the acceptance probability; trial moves are
made as usual. The direct integration of $\tilde{\beta}(E)$ is not desir-
able due to the sharp variation of the integrand at low
temperatures. To overcome this problem, we approxi-
mate the statistical temperature using the piecewise inter-
polation

$$\tilde{T}(E) = \tilde{T}_j + \lambda_j(E - E_j) \quad (4)$$

for $E_j \le E \le E_{j+1}$, where $\lambda_j = (\tilde{T}_{j+1} - \tilde{T}_j)/\Delta$ is the
slope of the linear segment connecting $[E_j, \tilde{T}_j]$ and
$[E_{j+1}, \tilde{T}_{j+1}]$. Linear extrapolation is particularly appropri-
ate at low temperatures, where the specific heat C_V is
nearly constant. The sequence of consecutive interpola-
tions also enables a faithful representation of a $T(E)$ cor-
responding to a phase transition [8]. The interpolation in
Eq. (4) yields the continuum entropy estimate,

$$\tilde{S}(E) = \int_{E_l}^{E} \tilde{\beta}(E')dE' = \sum_{j=l+1}^{i^*} L_j(E_j) + L_{i+1}(E), \quad (5)$$

where $i^* = i - 1$ (i) for $\bar{E}_{i-1} \le E \le E_i$ $(E_i \le E \le \bar{E}_i)$,
$\bar{E}_i = (E_i + E_{i+1})/2$, E_l is the lowest energy, and $L_j = \lambda_{j-1}^{-1} \ln[1 + \lambda_{j-1}(E - E_{j-1})/\tilde{T}_{j-1}]$. Equation (5) is used
to determine the acceptance of trial moves, taking into
account that the updated $\tilde{T}_j$ differs from its previous esti-
mate at only two grid points $i \pm 1$. For the corresponding
update of $\tilde{\Omega}$, let $\tilde{\Omega}_i = \exp[\tilde{S}(E_i)] = \prod_{j=l+1}^{i} Y_j$, $Y_j = [\tilde{T}_j/\tilde{T}_{j-1}]^{\lambda_{j-1}^{-1}}$. On a visit to E_i, Eq. (3) yields

$$\tilde{\Omega}'_{i+k} = \tilde{f}_k\tilde{\Omega}_{i+k}, \quad (6)$$

where $\tilde{f}_k = \Pi_{j=i-1}^{i+k} Y'_j/Y_j$ is a nonuniform modification
factor for $k \in [-1, 2]$; Y'_j is evaluated at the updated $\tilde{T}'_j$.

The combination of the fundamental Eq. (3) and the
intrinsic smoothing of Eq. (5) allows a continuum descrip-
tion of $\tilde{S}(E)$ regardless of the choice of Δ. Thus our method
can maintain statistical accuracy using larger values of Δ,
which is essential for large systems with a huge range of
$\Omega(E)$. In contrast, as seen below, Δ must be small for other
flat histogram MC methods [1,4].

Once the histogram fluctuations are less than 20% of the
mean, the sampling is repeated with a reduced convergence
factor $f_{n+1} = \sqrt{f_n}$, n being the iteration, and is terminated
at $f_d \equiv f - 1 < 10^{-8}$. During the initial stages of the
simulation, the temperature estimate for an unexplored
energy region is modified every 10^4 MC steps as $\tilde{T}(E) = T_{\min}$ for $E < E_{\min}$, $T_{\min} = \tilde{T}(E_{\min}) = \min\{\tilde{T}(E)\}$. This
low-energy flattening of $\tilde{T}(E)$ corresponds to the extrapo-
lation of $\tilde{S}(E) = \tilde{S}(E_{\min}) + (E - E_{\min})/T_{\min}$ for $E < E_{\min}$

and accelerates the convergence by assisting the system to
access lower energies through the canonical sampling at
$T_{\min}$. When the simulation converges with $f_d < 10^{-8}$,
thermodynamic properties are determined with the conver-
gent $\tilde{T}(E)$ via Eq. (5) [1].

We tested the STMC algorithm for the 32×32 Ising
model with periodic boundary conditions and $T_l = 1.2$,
$T_h = 4$, an initial constant temperature estimate $\tilde{T}(E) = T_h$, and considered $\Delta = 8, 32, 64$, and an initial value
$f_{d,0} = 10^{-5}$. Note that f_0 is very close to unity due to
the restricted sampling range of $\tilde{T}(E)$, in contrast to WL
sampling, which usually starts with $f_0 = e$ to cover a large
range of $\tilde{\Omega}(E)$. Consequently, both $\tilde{T}(E)$ and $\tilde{S}(E)$ are
almost indistinguishable after the first iteration from their
convergent values with $f_d = 10^{-8}$. When Δ is increased to
64, the temperature estimate displays a small ruggedness
due to the discrete nature of the histogram [Fig. 1(a)], but
the flat histograms in the inset confirm that STMC works
even for large energy bins. The corresponding entropy
estimates, Fig. 1(b), show good agreement with the exact
result [9], and the errors $\delta\tilde{S}(E) = \tilde{S}(E) - S(E)$ are less
than one for the region $[E(T_l), E(T_h)]$, regardless of the
value of Δ [inset, Fig. 1(b)].

Another test is the continuum $N = 110$ Lennard-Jones
fluid with a cutoff of 2.5σ at reduced density $\rho = 0.88$,
$T_l = 0.65$, and $T_h = 1.82$. The flatness of the histogram
has been checked every 10^4 steps for the energy window
$-670 \le E \le -540$. We found [see Fig. 2(a)] a small
ruggedness in $\tilde{T}(E)$ with E for $\Delta = 1$, $f_{d,0} = 10^{-3}$, but
$\tilde{T}(E)$ shows a smoother variation with $f_{d,0} = 10^{-4}$ and
$\Delta = 4$ and 16. Nevertheless, the simulations give flat histo-
grams [inset of Fig. 2(a)] for all values of Δ. To demon-
strate the progression of statistical errors we have checked
the average standard deviations of the specific heat C_V of
five independent runs at temperatures $T = 1.0, 1.1, 1.2,$
1.3, and 1.4 [see Fig. 2(b)]. Precise estimates of C_V were
determined by canonical sampling for 10^6 steps. The sta-
tistical errors of both STMC and WL rapidly decrease to a
limiting value because, as f_d decreases, further MC steps
do not [3] refine $\tilde{T}(E)$ or $\tilde{S}(E)$. However, STMC is superior
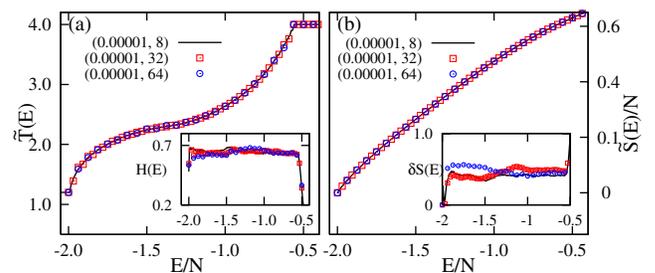to WL for the same $\Delta = 1$ and is comparable even for

FIG. 1 (color online).   The 32×32 Ising model: (a) convergent
$\tilde{T}(E)$ and histograms (inset); (b) entropy estimates $\tilde{S}(E)$ and
absolute errors $\delta S(E)$ (inset) for energy bins $\Delta = 8, 32,$ and
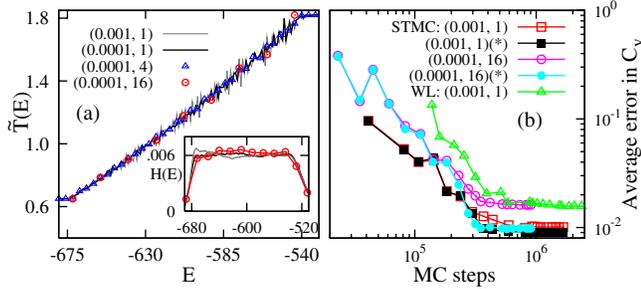64. The notation for curves is $(f_{d,0}\Delta)$.

FIG. 2 (color online). The $N = 110$ LJ fluid: (a) convergent $\tilde{T}(E)$ and histograms (inset); (b) statistical errors of STMC and WL in specific heats as a function of MC steps with various values of $(f_{d,0}, \Delta)$. Reweighted results indicated by * have been obtained from the initial simulation data of 1.5×10^6 MC steps with $\Delta = 1$ and 10^6 MC steps with $\Delta = 16$.

$\Delta = 16$ with fewer MC steps. The advantage is due to the continuum description of $\tilde{S}(E)$, because STMC calculates an entropy difference for trial moves in which E' and E both belong to the same energy bin, while WL always accepts such moves. Accuracy can be further improved with reweighting [10], which corrects the entropy as $S(E) = \tilde{S}(E) + \ln P(E)$, $P(E)$ being the normalized energy histogram. The improvement increases with bin size, as does the rate of convergence [Fig. 2(b)], so at $\Delta = 16$ STMC has accelerated convergence and almost the same accuracy as with $\Delta = 1$. By contrast, we found that the convergence of the WL algorithm slows significantly with increasing Δ in both the fluid and the Ising model, precluding the option of using a large bin and reweighting.

Despite the advantage of using larger values of Δ demonstrated by STMC, the formulation of STMD is our primary result. The effective potential $T_0\tilde{S}(E)$ is employed, and the velocity distribution is maintained at the temperature T_0 with a Nose-Hoover thermostat [11]. The force $\tilde{\mathbf{f}}_i$ on particle i is related to the true force $\mathbf{f}_i$ by the energy dependent scaling, $\gamma(E) = T_0/\tilde{T}(E)$,

$$\tilde{\mathbf{f}}_i = \gamma(E)\mathbf{f}_i, \qquad (7)$$

where E denotes the potential energy, and the statistical weight for configurations in an MD trajectory is $e^{-\tilde{S}(E)} = 1/\tilde{\Omega}(E)$, corresponding [5] to a flat energy distribution with $S(E) = \tilde{S}(E)$. We convert force-scaled MD to STMD by updating the statistical temperature every time step, using the same procedure as in STMC.

We studied STMD for the off-lattice coarse-grained BLN protein model [12], composed of hydrophobic (B), hydrophilic (L), and neutral (N) beads, and including bond stretches, bends, torsions, and realistic nonbonded interactions. The 46-mer sequence, $B_9N_3(LB)_4N_3B_9N_3(LB)_5L$, and the 69-mer sequence, $B_9N_3(LB)_4N_3B_9N_3(LB)_4N_3B_9N_3(LB)_5L$, which exhibit 4 and 6-stranded β-barrel global potential energy minima (native states), were tested with the same potential and

TABLE I. Number of inherent structures with energy less than ΔE above the global minimum for BLN 46-mer and 69-mer found by STMD, conformational space annealing (CAS) [14], and automated histogram filtering (AHF) [15].

	46-mer					69-mer			
ΔE	1	2	3	5	ΔE	1	2	3	5
STMD	5	40	189	1045	STMD	3	44	205	1389
CAS	5	36	147	636	AHF	3	47	175	935

parameter set [13]. Because of the ruggedness of the potential energy landscape, the 46-mer has been used as a benchmark for global energy minimization [13,14], and the increased complexity of the 69-mer presents a more stringent test for our algorithm [15]. Simulations were started with $f_{d,0} = 1.0005/\Delta$ with $T_l = 0.1$ and $T_h = T_0 = 1.3$, and $\Delta = 1$ and 2 for the 46- and 69-mer, respectively. The known global minima of the 46-mer and 69-mer are reproduced at $E_0 = -49.2635$ [13] and -99.189 after 10^7 and 2.8×10^7 MD steps, respectively [15]. A comparison of the numbers of local energy minima [inherent structures [16]] found by STMD with those found by sophisticated optimization algorithms [14,15] in Table I illustrates the ability of STMD to find low-energy states, overcoming a broken ergodicity.

A previous study of the 46-mer [17] has shown the existence of characteristic collapse and folding transitions as reflected in the specific heat C_V. As expected, the collapse transition is associated with a peak in C_V [Fig. 3(a)]. However, the secondary peak, which has been interpreted as the signature of folding at $T \sim 0.34$ [17], is significantly reduced and appears only as a small shoulder in $\partial C_V/\partial T$. This suppression can be understood in terms of the canonical average occupation probability, $p_i(T)$ [Fig. 3(b)], of the configuration-space basin of the ith inherent structure, obtained by reweighting the p_i from the STMD trajectory using $\tilde{S}(E)$. Folding involves occupation of the lowest-lying basin with $i = 0$. We find that individual p_i become nonzero below the collapse tempera-
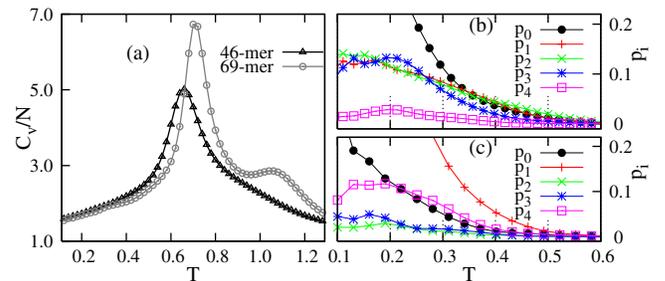

FIG. 3 (color online). (a) Specific heat C_V for BLN 46-mer and 69-mer, with peaks corresponding to the collapse transition. The mean inherent structure occupation probability $p_i(T)$ for the ith inherent structure for (b) 46-mer and (c) 69-mer (p_0 corresponds to the global minimum).

TABLE II.   Global energy minima of 55-mer AB models with a Fibonacci sequence in two and three dimensions determined by STMD, the pruned-enriched-Rosenbluth method with importance sampling (nPERMis) [19], annealing contour MC (ACMC) [20], the energy landscape paving (ELP) [21], and conformational space annealing (CSA) [22].

Dimension	nPERMis	ACMC	ELP	CSA	STMD
2	−18.5154	−18.7407		−18.9110	−18.9202
3	−32.8843		−42.428	−42.3418	−42.5781

ture, and those for higher energy inherent structures are still nonzero down to $T \sim 0.1$; their contributions reduce the thermodynamic signature of folding. The point is that the smoothing of $\tilde{T}(E)$ occurs only if the occupation of these states is accurately sampled. Our result that $p_{i=1,2} > p_0$ after collapse reveals why global optimization of the 46-mer so often fails to find the global minimum [13,14].

Except for a local ordering transition around $T \sim 1.05$ in Fig. 3(a), thermodynamic behaviors of collapse and folding in the 69-mer are similar to those of the 46-mer. The profile of $p_i(T)$ in Fig. 3(c) reflects the increased roughness of the potential energy landscape. The excited state occupation probabilities p_1 and p_4 are much greater than p_0 around the collapse temperature and p_1 is still higher than p_0 up to $T \sim 0.1$. This increased accessibility to non-native inherent structures during the collapse leads to a substantial slowing down of folding through kinetic trapping in misfolded excited states.

Because STMD yields a flat energy distribution, it can find global minima in complex systems where effective MC moves are not feasible. As a test, we considered two- and three-dimensional off-lattice AB protein models [18]

consisting of two types of hydrophobic (A) and hydrophilic (B) monomers, which have been the subject of various MC optimization algorithms [19–22]. For a very small $T_l = 0.02$ and $f_{d,0} = 0.0005$, we studied four different size chains S_{13}, S_{21}, S_{34}, and S_{55} in two and three dimensions. Configurations associated with $\tilde{T}(E)/T_l \leq 1.5$ were minimized to generate inherent structures. Our method reproduced all the known ground states for chain lengths ≤ 34 [22] and found new candidates for global minima in S_{55} for both dimensions (Table II and Fig. 4), confirming the promise of STMD-based global minimization.

*Electronic address: jaegil@bu.edu

[1] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001); Phys. Rev. E **64**, 056101 (2001).
[2] Q. Yan and J. J. Pablo, Phys. Rev. Lett. **90**, 035701 (2003).
[3] M. S. Shell, P. G. Debenedetti, and A. Z. Panagiotopoulos, J. Chem. Phys. **119**, 9406 (2003).
[4] B. A. Berg and T. Celik, Phys. Rev. Lett. **69**, 2292 (1992); J. Lee, Phys. Rev. Lett. **71**, 211 (1993).
[5] N. Nakajima, H. Nakamura, and A. Kidera, J. Phys. Chem. B **101**, 817 (1997); U. H. E. Hansmann, Y. Okamoto, and F. Eisenmenger, Chem. Phys. Lett. **259**, 321 (1996).
[6] J. G. Kim, Y. Fukunishi, and H. Nakamura, Phys. Rev. E **67**, 011105 (2003).
[7] B. J. Schulz, K. Binder, M. Müller, and D. P. Landau, Phys. Rev. E **67**, 067102 (2003).
[8] J. G. Kim, Y. Fukunishi, and H. Nakamura, J. Chem. Phys. **121**, 1626 (2004).
[9] P. D. Beale, Phys. Rev. Lett. **76**, 78 (1996).
[10] U. H. E. Hansmann and Y. Okamoto, Phys. Rev. E **54**, 5863 (1996).
[11] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).
[12] J. D. Honeycutt and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **87**, 3526 (1990).
[13] Y.-H. Lee and B. J. Berne, J. Phys. Chem. A **104**, 86 (2000).
[14] S. Y. Kim, S. J. Lee, and J. Lee, J. Chem. Phys. **119**, 10 274 (2003).
[15] S. A. Larrass, L. M. Pegram, H. L. Gordon, and S. M. Rothstein, J. Chem. Phys. **119**, 13 149 (2003).
[16] F. H. Stillinger and T. A. Weber, Phys. Rev. A **28**, 2408 (1983).
[17] Z. Guo and C. L. Brooks III, Biopolymers **36**, 83 (1995); **42**, 745 (1997).
[18] F. H. Stillinger, Phys. Rev. E **48**, 1469 (1993).
[19] H.-P. Hsu, V. Mehra, W. Nadler, and P. Grassberger, J. Chem. Phys. **118**, 444 (2003).
[20] F. Liang, J. Chem. Phys. **120**, 6756 (2004).
[21] M. Bachmann, H. Arkin, and W. Janke, Phys. Rev. E **71**, 031906 (2005).
[22] S. Kim, S. B. Lee, and J. Lee, Phys. Rev. E **72**, 011916 (2005).
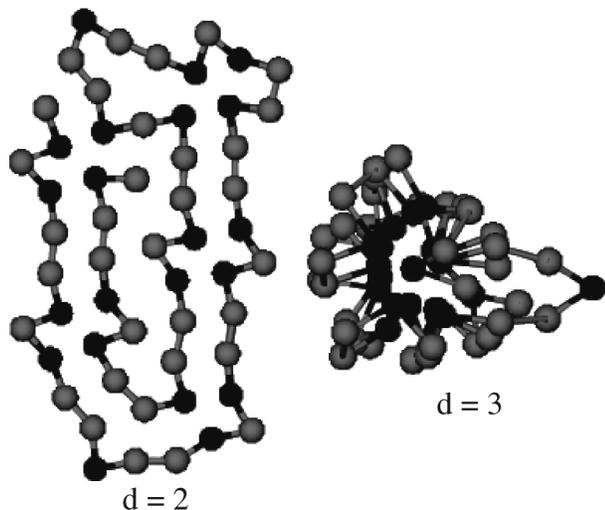

FIG. 4.   Structures of new global minima of S_{55} in $d = 2$ and 3 dimensions. Gray and black balls correspond to hydrophilic and hydrophobic monomers, respectively.