Available online at www.sciencedirect.com
SCIENCE DIRECT®
Chemical Physics Letters 410 (2005) 430–435
CHEMICAL PHYSICS LETTERS
www.elsevier.com/locate/cplett

Calculation of the entropy of lattice polymer models from Monte Carlo trajectories

Ronald P. White, Jason Funt [1], Hagai Meirovitch *

Department of Computational Biology, University of Pittsburgh School of Medicine, Biomedical Science Tower, 1058w, Pittsburgh, PA 15261, United States

Received 11 May 2005

Abstract

While lattice models are used extensively for macromolecules (synthetic polymers proteins, etc.), calculation of the *absolute* entropy, S, and the free energy, F, from a given Monte Carlo (MC) trajectory is not straightforward. Recently, we have developed the hypothetical scanning MC (HSMC) method for calculating S and F of fluids. Here we extend HSMC to self-avoiding walks on a square lattice and discuss its wide applicability to complex polymer lattice models. HSMC is independent of existing techniques and thus constitutes an independent research tool; it provides rigorous upper and lower bounds for F, which can be obtained from a very small sample and even from a *single* chain conformation.

Lattice models have been utilized to study a wide range of phenomena in polymer physics [1–5] as well as in structural biology, mainly related to protein folding and stability [6–9]. (Refs. [1–9] constitute a very limited representation of hundreds of papers published in the last 15 years.) Because of their simplicity these models have been invaluable tools for understanding global properties that do not depend strongly on molecular details. Such models vary in complexity, ranging from self-avoiding walks on a square lattice to chain models on enriched 3D lattices with a large effective coordination number.

Commonly, these systems are simulated by variants of Metropolis Monte Carlo (MC) – a dynamical method that enables one to generate samples of chain configurations i distributed according to their Boltzmann probability, P_i^{B}, from which equilibrium information can be extracted [10]. Using MC it is straightforward to calculate properties that are measured directly from i, such as the potential energy E_i. On the other hand, the *value* of P_i^{B} cannot be obtained in a straightforward manner, which makes it difficult to calculate the *absolute* entropy, $S \sim -\ln P_i^{\mathrm{B}}$ *directly*, i.e., as a byproduct of the simulation (like E_i). There is a strong interest in S as a measure of order and as an essential ingredient of the free energy, $F = E - TS$, where T is the absolute temperature; F constitutes the criterion of stability, which is mandatory in structure determination of proteins, for example. Furthermore, because MC simulations constitute models for dynamical processes, one would seek to calculate changes in F and S during a relaxation process, by assuming local equilibrium in certain parts along the MC trajectory; a classic example is simulation of protein folding [11].

S and F are commonly calculated by thermodynamic integration (TI) techniques [12–14] that do not operate on a given MC sample but requires conducting a *separate set* of MC simulations. This is a robust approach that enables one to calculate differences, $\Delta S_{\mathbf{ab}}$ and $\Delta F_{\mathbf{ab}}$, between two states **a** and **b** of a system; however, if the

* Corresponding author. Fax: +1 412 648 3163.
E-mail address: hagaim@pitt.edu (H. Meirovitch).
[1] A BBSI student, summer 2004.

0009-2614/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.cplett.2005.06.002

structural variance of such states is large (e.g., helical and hairpin states of a polypeptide) the integration from state **a** to **b** becomes difficult and in many cases unfeasible. On the other hand, if one could calculate the absolute $F_{\mathbf{a}}$ and $F_{\mathbf{b}}$ directly from two separate sets of simulations carried out at states **a** and **b**, $\Delta F_{\mathbf{ab}} = F_{\mathbf{a}} - F_{\mathbf{b}}$ and the integration can be avoided. Still, the absolute F can also be obtained with TI provided that a reference state **r** is available, where the free energy is known exactly and the integration path between **r** and **a** (and **b**) is relatively short. However, for non-homogeneous lattice models such integration might not be trivial, and in models of peptides and proteins, defining reference states that are close to the state of interest is a standing problem.

Another type of simulation method has been developed for polymers, where a chain is constructed step-by-step with transition probabilities (TPs) ([15–19], see also an extensive review in [5]). The product of these TPs leads to P_i^{B}, hence S is known. However, these build-up procedures are not always the methods of choice mainly because they lack the dynamical aspects (and simplicity) of MC, which thus has become the commonly used method. Hence, it is important to develop methods for calculating the absolute entropy from a given MC trajectory. Nonetheless, a hybrid of one build-up procedure, the scanning method [19], with the dynamical MC approach has led to two *approximate* techniques, the local states (LS) [20,21] and hypothetical scanning (HS) methods [22,23]. These methods enable one to calculate S and F directly from a *given* sample generated by *any* simulation technique, and they have been applied successfully to polymers, peptides, proteins, magnetic systems, and lattice gas models [14].

Recently, the HS method has been extended to fluids and has been further developed by defining TPs that are calculated by an MC procedure and (unlike the TPs of HS) take into account *all* the long-range interactions [24,25]; this HSMC method has been applied very successfully to liquid argon, TIP3P water [25], and polyglycine molecules in helical, extended and hairpin states [26]. HSMC is significantly more accurate than HS, provides rigorous upper and lower bounds for F, which can be calculated from a relatively small sample and even from a *single* conformation.

The aim of this paper is to extend the scope of HSMC to lattice polymer models, in particular to random coil chains. For that we study self-avoiding walks (SAWs) on a square lattice – a difficult test case due to the strong excluded volume (EV) interactions occurring in 2D [5] – and discuss application of HSMC to more complex lattice chain systems. The present results are compared to results obtained by us using TI, to those obtained some time ago by the scanning method [27], and to results based on series expansion (exact enumeration) techniques [28]. In what follows we first describe the scanning method [19], the HS method, and then HSMC for SAWs.

Assume a single SAW of N steps (bonds), i.e., $N+1$ monomers starting from the origin on a square lattice. All the SAWs i are equally probable with Boltzmann probability

$$P_i^{\mathrm{B}} = 1/Z_{\mathrm{SAW}}, \tag{1}$$

where the partition function, Z_{SAW}, is the total number of different SAWs, and the free energy is

$$F/k_{\mathrm{B}}T = -S/k_{\mathrm{B}} = \sum_i P_i^{\mathrm{B}} \ln P_i^{\mathrm{B}} = -\ln Z_{\mathrm{SAW}} = \ln P_j^{\mathrm{B}}, \tag{2}$$

where k_{B} is the Boltzmann constant and j is *any* SAW. The summations (in i) here and in the rest of the paper are over the *ensemble* of SAWs. Eq. (2) demonstrates that F (and S for this particular model) has zero fluctuation, which is a general property of the *correct* free energy of any system, while the fluctuation of an *approximate* F is expected to be finite [29]. Eq. (2) also shows that if the Boltzmann probability of any single SAW (j) is known, F (and S for this particular model) is known as well, which again is a general property satisfied by any system in equilibrium.

With the scanning method [19] a SAW is grown step-by-step with TPs; thus, at step k of the process, $k-1$ directions (bonds), v ($v = 1,4$) will have already been constructed [they are denoted $v_1,\ldots,v_{(k-1)}$]. To determine the direction v_k (out of 4 possible directions, v) one enumerates all the possible continuations $Z_k^v(f)$ of the chain in a limited number of f future steps that start from v of step k, where $Z_k^v(f)$ is a partial future partition function and f is the scanning parameter. $Z_k^v(f)$ enables one to define TPs for v,

$$p(v|v_{(k-1)},\ldots,v_1,f) = Z_k^v(f) \Big/ \sum_{v=1}^{4} Z_k^v(f). \tag{3}$$

Using these TPs, the kth step is determined by a random number and the process continues. The construction probability $P_i^0(f)$ of SAW i is the product of the TPs with which the steps have been chosen,

$$P_i^0(f) = \prod_{k=1}^{N} p(v_k|v_{(k-1)},\ldots,v_1,f). \tag{4}$$

Again, for $f \ll N$, $P_i^0(f)$ is approximate. Due to this 'incomplete' scanning, the chain can get trapped in a dead end during construction. Also, $P_i^0(f)$ is biased, i.e., unlike P_i^{B}, it is larger for the compact SAWs than for the open ones. This bias can be decreased *systematically* by increasing f, where for a complete future scanning, i.e., $f_{\max} = N - k + 1$, the TPs (Eq. (1)) become exact and no trapping occurs [19]. In practical applications the bias is removed by an *importance sampling* procedure, which leads to an unbiased estimation that is

exact within the statistical error. The scanning method can easily be extended to a chain model with finite interactions; in this case the interaction energy $E^k_{j(v)}(f)$ of the future chain j that starts from v with itself and with the rest of the chain is calculated and the corresponding Boltzmann factor contributes to $Z^v_k(f)$, rather than 1, $Z^v_k(f) = \sum_{j(v)} \exp[-E^k_{j(v)}(f)/k_B T]$.

The HS method (as well as LS) is based on the concept that two samples in equilibrium generated by different simulation methods are equivalent in the sense that both lead to the same estimates (within the statistical error) of average properties, such as the entropy, energy, and their fluctuations. Relying on this equivalence, one assumes that a given sample of SAWs constructed by *any* exact procedure (e.g., Metropolis MC) has instead been generated with the scanning method. Thus, for each of the bonds $[v_k(i)]$ of SAW i one calculates the TPs (Eq. (3)) as if i had been generated with the scanning method. The product of these TPs leads to $P^0_i(f)$ (Eq. (4)) and to a functional S^A, which can be shown *rigorously* (using Jensen's inequality) to be an upper bound for S [23],

$$S^A(f) = -k_B \sum_i P^B_i \ln P^0_i(f), \tag{5}$$

where i runs on the *complete* ensemble of SAWs. The fluctuation $\sigma_A(f)$ of $\ln P^0_i(f)$,

$$\sigma_A(f) = \left\{ \sum_{\text{SAWs } i} P^B_i [S^A(f) + k_B \ln P^0_i(f)]^2 \right\}^{1/2}, \tag{6}$$

is expected to be larger than zero, decreasing with increasing f (i.e., with improving the approximation).

While the TPs defined by HS are deterministic (based on *all* the future SAWs of f bonds at step k), for a large chain they are always approximate, i.e., $f \ll N$ due to the exponential growth (with f) of the number of future SAWs. The HSMC method overcomes this limitation by seeking to estimate the *exact* TP at step k (see Eq. (3)),

$$p(v|v_{(k-1)}, \ldots, v_1, f_{\max} = N - k + 1) = Z^v_k(f_{\max}) \Big/ \sum_{v=1}^{4} Z^v_k(f_{\max}). \tag{7}$$

Thus, an MC simulation of the *entire* future part of the chain (i.e., steps $k, k+1, \ldots, N$) is carried out in the presence of the 'frozen past' $[v_1, \ldots, v_{(k-1)}]$. The TP of the actual direction, $v_k(i)$ in the reconstructed SAW i is calculated from the number of MC steps, $n^{v(i)}_k$ for which $v_k(i)$ was visited during the simulation of total n_{MC} MC steps at k,

$$p^{HS}(v_k(i)|v_{(k-1)}, \ldots, v_1) = n^{v(i)}_k / n_{MC} \tag{8}$$

and the reconstruction probability of chain i is

$$P^{HS}_i = \prod_{k=1}^{N} p^{HS}(v_k|v_{(k-1)}, \ldots, v_1), \tag{9}$$

where, for simplicity, i has been omitted in the TPs. To be consistent with [25], the probabilities, P^{HS}_i and p^{HS}, are superscripted with HS rather than HSMC. It should be noted that unlike the deterministic $P^0_i(f)$ (Eq. (4)), P^{HS}_i is defined stochasticly. The fact that the entire future is considered is important for systems with strong long-range interactions such as SAWs, proteins, etc. Still, p^{HS} and hence P^{HS}_i are approximate, but as the MC simulation is increased, their estimation improves, i.e., $p^{HS} \to p^{exact}$ and $P^{HS}_i \to P^B_i$, meaning that S can be estimated by reconstructing a *single* SAW (see Eq. (2)). In practice, however, P^{HS}_i is approximate leading to an approximate functional S^A (compare with Eq. (5))

$$S^A = -k_B \sum_i P^B_i \ln P^{HS}_i = \sum_i P^B_i S^{HS}_i. \tag{10}$$

It can be shown (see Appendix of [25]) that like S^A (Eq. (5)), S^A (Eq. (10)) defined with stochastic probabilities, P^{HS}_i, is a rigorous upper bound, which is expected to have non-zero fluctuation σ_A (Eq. (6)). Also, it should be pointed out that an HSMC reconstruction for SAWs with attractions is practically the same, where, however, the MC acceptance criterion is determined by both, EV and the attractions [26].

One can define another entropy functional, S^B that is a rigorous lower bound of S. To estimate S^B from an (exact) MC sample, we express it in terms of statistical averages defined with P^B_i,

$$S^B = -k_B \sum_i P^{HS}_i \ln P^{HS}_i = -k_B \frac{\sum_i P^B_i [P^{HS}_i \ln P^{HS}_i]}{\sum_i P^B_i P^{HS}_i}. \tag{11}$$

If the deviations of S^A and S^B from S (in the absolute values) are approximately equal, their average S^M becomes a better approximation than either of them individually,

$$S^M = [S^A + S^B]/2. \tag{12}$$

The entropy can be expressed *exactly* by S^D (see [25]), which can also be estimated from a sample generated with P^B_i

$$S^D = -k_B \ln \sum_i P^B_i P^{HS}_i = -k_B \ln \sum_i P^B_i [\exp(-S^{HS}_i/k_B)]. \tag{13}$$

While the theory above has been introduced for the entire ensemble, it also applies to a set of reconstructions of a single chain conformation (see Appendix of [25]). Thus, we have calculated the entropy of SAWs consisting of $N = 49$, 99, 149, 249, 399 and 599 bonds, where for each chain length the results were obtained by n replicate reconstructions (based on a different sets of random numbers) of a *straight* SAW of N bonds. For example, in this paper S^A is estimated as follows: from n reconstructions of the same single chain we obtain n values for $\ln P^{HS}_t$ and we take their arithmetic average $-(k_B/n)\sum_t \ln P^{HS}_t$; an analogous procedure is used

for S^B and S^D. The efficiency of HSMC is affected considerably by the MC procedure employed in the reconstruction process. On a square lattice, 'crankshaft' moves are in most cases rejected due to the strong EV interactions while corner moves have somewhat higher acceptance rate [5]. Therefore, for the reconstruction process we have used an MC procedure based on 50% corner moves (that provide local conformational changes) and 50% 'pivot' moves that have been shown to effectively induce global changes [30].

The calculations are based on the sample size n – the number of reconstructed SAWs and n_{future}, which is related to the number of future MC steps per bond applied during the reconstruction process as defined below. First we note that the first bond of the chain is not reconstructed; its probability is always 1/4. The number of MC steps, n_{MC}, for bond k is scaled as $n_{\text{MC}} = (N-k+1)n_{\text{future}}$, meaning that the maximal number of future MC steps is applied for the reconstruction of the second bond (to which corresponds the largest future segment of $N-1$ bonds), while the last bond (N) is allotted the minimal number of MC steps. Because each simulation at step k always starts from a straight chain it is important to let the future SAW equilibrate, otherwise p^{HS} (Eq. (8)) would (on average) be too high; therefore, 300 MC steps per future bond are used for equilibration. As discussed earlier, the larger is n_{future} the better (i.e., smaller) is S^A (Eq. (10)), the larger is S^B (Eq. (11)) and the smaller is the fluctuation, σ_A (Eq. (6)). To demonstrate this effect, the results for each chain length are presented in Table 1 for $n_{\text{future}} = 500$, 5000, and 50 000, where the corresponding sample size, n, is decreased, which results in approximately the same computer time for each calculation. We present results obtained with the scanning method [27] and with series expansion $[S/k_B = (\ln c_N)/N]$, where $c_N \sim \mu^N[a_1 N^{11/32} + a_2\ N^{-21/32} + b_1 N^{-37/32} + (-1)^N d_1 N^{-3/2} + (-1)^N\ d_2 N^{-2}]$, $a_1 = 1.1771(2)$, $a_2 = 0.554(2)$, $b_1 = -0.19(2)$, $d_1 = -0.19(2)$, $d_2 = 0.034(2)$, and $\mu = 2.6381585(10)$ (the error of the last digit appears in parenthesis) [28]. Also, using the present MC procedure, we have carried out TI simulations starting from an ideal chain (with known entropy of $k_B[\ln 4 + (N-1)\ln 3]$) and integrating S by a gradual increase of the EV interaction (e.g., see [31]). These results are presented in the table as well. We shall consider the TI and series results as correct.

The table supports the above expectations. Thus, for all chain lengths, as n_{future} is increased from 500 to 50 000, the fluctuation decreases, S^A decreases and remains an upper bound, and S^B increases remaining a lower bound. On the other hand, for $N \leqslant 249$, S^M the average of S^A and S^B is constant for the three n_{future} values and for $N = 399$ and 599 S^M is the same for $n_{\text{future}} = 5000$ and 50 000. In all these cases S^M is equal, within the error bars, to the TI and series results, and for $N \ll 599$ also to the scanning results, which demonstrates that for these cases (i.e., for good enough approximations) the absolute values of S^A and S^B deviate equally from the correct results. For each N, S^D and S^M are equal within the statistical errors. We suspect that the scanning result for $N = 599$ underestimates the correct value due to the bias (toward the compact SAWs) introduced by the scanning procedure, which has not been removed completely by importance sampling. Also, the series expansion and TI results are equal within the error bars except for $N = 599$. Overall the HSMC statistical errors are small (0.002–0.005%); however, it should be noted that much more computer time has been invested in the simulations of the longer chains.

An inherent inefficiency of HSMC lies in the need to carry out $N-1$ simulations for an N-bond SAW. Still, performance can be improved by changing the scaling function discussed above, which controls the extent of simulation applied to each bond in the reconstruction process. However, the most significant factor affecting efficiency is the simulation method used. Thus, our preliminary simulations based on corner moves alone have converged extremely slowly, and adding the pivot moves improved performance dramatically. In three dimensions, where the EV effect is weaker, one can add crankshaft moves (and other moves, see [5]) that are expected to increase efficiency further. Also, a chain with attractive interactions (a homopolymer or a heteropolymer consisting of monomers with different interactions) unlike SAWs would span (at low T) only a limited part of conformational space; to obtain the corresponding *local F*, the future chains should be limited to this region, which can be achieved only by local MC moves [26]. Moreover, in this work we have studied straight chains that are the easiest to reconstruct, where in practical applications non-straight SAWs will be treated. For such chains one can envisage situations where the present MC procedure will not be ergodic (at least for specific bonds) due to geometrical constraints imposed by the frozen past, thus leading to incorrect probabilities ρ^{HS} (Eq. (8)). One remedy for this problem would be to replace for these bonds the present *dynamic* MC procedure by a suitable step-by-step construction (growth) procedure [15–18] (these procedures can provide S, but unlike HSMC, not from a given trajectory). For SAWs the most efficient is the scanning method, followed by TI, where HSMC is the least efficient. For example, one reconstruction of a 399-bond SAW for $n_{\text{future}} = 50\,000$ requires ~4.2 h CPU leading to $S = 0.9757$ (6). The value of TI in the table required ~100 h CPU.

However, the applicability of HSMC to both random coil SAWs and peptides that fluctuate locally [26] demonstrates applicability to all ranges of flexibility, versatility that is not shared by other methods. Thus, the harmonic and quasi-harmonic techniques [32,33] are limited to handle (at least approximately) local fluctuations (for which HS has failed), LS is very inefficient

Table 1
HSMC results for the entropy of N-bond SAWs obtained from n reconstructions of a straight chain

n_{future}	S^A/k_B	σ_A	S^B/k_B	S^M/k_B	S^D/k_B	n
$N = 49$, $S_{SCAN} = 1.000904(4)$						
500	1.00583 (1)	0.01424 (2)	0.99602 (5)	1.00093 (3)	1.00091 (3)	1 250 000
5000	1.00140 (1)	0.00448 (2)	1.00042 (3)	1.00091 (2)	1.00091 (2)	125 000
50 000	1.00095 (1)	0.00142 (2)	1.00085 (3)	1.00090 (2)	1.00090 (2)	12 500
S_{TI}	1.000897 (3)		1.000897 (3)	1.000897 (3)	1.000897 (3)	
Sseries	1.000899 (4)		1.000899 (4)	1.000899 (4)	1.000899 (4)	
$N = 99$, $S_{SCAN} = 0.987726(5)$						
500	0.99294 (2)	0.01030 (3)	0.9826 (1)	0.98775 (5)	0.98773 (5)	250 000
5000	0.98826 (2)	0.00324 (3)	0.98722 (5)	0.98774 (3)	0.98774 (3)	25 000
50 000	0.98777 (2)	0.00101 (3)	0.98767 (4)	0.98772 (2)	0.98772 (3)	2500
S_{TI}	0.987727 (3)		0.987727 (3)	0.987727 (3)	0.987727 (3)	
Sseries	0.987730 (3)		0.987730 (3)	0.987730 (3)	0.987730 (3)	
$N = 149$, $S_{SCAN} = 0.982740(3)$						
500	0.98806 (2)	0.00852 (3)	0.9774 (2)	0.9827 (1)	0.9827 (1)	250 000
5000	0.98329 (2)	0.00267 (3)	0.98222 (5)	0.98276 (3)	0.98276 (3)	25 000
50 000	0.98281 (2)	0.00085 (3)	0.98270 (4)	0.98275 (2)	0.98275 (3)	2500
S_{TI}	0.982742 (3)		0.982742 (3)	0.982742 (3)	0.982742 (3)	
Sseries	0.982740 (2)		0.982740 (2)	0.982740 (2)	0.982740 (2)	
$N = 249$, $S_{SCAN} = 0.97836(2)$						
500	0.98391 (3)	0.00669 (4)	0.9727 (3)	0.9783 (2)	0.9783 (2)	63000
5000	0.97889 (2)	0.00208 (4)	0.97782 (8)	0.97836 (4)	0.97836 (5)	9100
50 000	0.97840 (2)	0.00066 (4)	0.97829 (5)	0.97835 (3)	0.97835 (3)	930
S_{TI}	0.978358 (4)		0.978358 (4)	0.978358 (4)	0.978358 (4)	
Sseries	0.978360 (1)		0.978360 (1)	0.978360 (1)	0.978360 (1)	
$N = 399$, $S_{SCAN} = 0.97567(4)$						
500	0.98138 (6)	0.00540 (5)	0.9710 (5)	0.9762 (3)	0.9759 (3)	9500
5000	0.97625 (4)	0.00170 (5)	0.9751 (1)	0.97567 (5)	0.97567 (5)	2000
50 000	0.97568 (4)	0.00053 (5)	0.97557 (7)	0.97563 (4)	0.97563 (5)	225
S_{TI}	0.975655 (8)		0.975655 (8)	0.975655 (8)	0.975655 (8)	
Sseries	0.975652(1)		0.975652 (1)	0.975652 (1)	0.975652 (1)	
$N = 599$, $S_{SCAN} = 0.97395(5)$						
500	0.98003 (8)	0.00445 (7)	0.9706 (8)	0.9753 (4)	0.9748 (5)	3000
5000	0.97466 (7)	0.00139 (7)	0.9736 (2)	0.9741 (1)	0.9741 (1)	450
50 000	0.97413 (5)	0.00036 (7)	0.9741 (1)	0.97409 (6)	0.97409 (5)	45
S_{TI}	0.97404 (1)		0.97404 (1)	0.97404 (1)	0.97404 (1)	
Sseries	0.974025(1)		0.974025(1)	0.974025(1)	0.974025(1)	

S^A (Eq. (10)) and S^B (Eq. (11)) are upper and lower bounds, respectively, S^M (Eq. (12)) is their average, and S^D (Eq. (13)) is an exact entropy functional. σ_A (Eq. (6)) is the fluctuation and n_{future} is related to the number of MC steps per bond (see text). S_{TI}, S_{scan}, and S_{series} were obtained by thermodynamic integration, the scanning method [27], and a series expansion formula (see text), respectively. The statistical error is defined by parentheses: $1.00(3) = 1.00 \pm 0.03$.

for SAWs, and calculating the *absolute* S (and F) of local fluctuations of peptides by TI is a standing problem. The practical application of HSMC to a wide range of lattice models (e.g., with attractions or any set of boundary conditions) is straightforward but requires selecting an optimal simulation method for each case, as discussed earlier. An interesting test case is a model of multiple SAWs enclosed in a 'box', studied previously by the scanning and HS methods [34], where chains are added successively to an initially empty box. However, with HS only the *partial* future of a reconstructed chain is considered, whereas HSMC can take into account the entire future, including that of the reconstructed chain and the positions and conformations of the as yet unreconstructed chains.

In summary, calculation of S is a central (notoriously difficult) problem in computer simulation and HSMC with its unique features constitutes a new tool for obtaining S independent of other methods. With HSMC all interactions are considered, and its accuracy depends only on the amount of MC sampling. Furthermore, the accuracy analysis of the results (S^M and S^D) is inherent in the method, by verifying the increase and decrease of the rigorous upper and lower bounds, S^B and S^A, and the decrease of σ_A, as the approximation improves. Finally, HSMC is of general applicability and unlike most methods, enables one to extract the absolute entropy from a given sample, where only a small number of SAWs (and even a single chain) need to be reconstructed; this is impor-

tant for studying relaxation processes, such as protein folding.

Acknowledgments

This work was supported by NIH Grants R01 GM61916 and R01 GM66090.

References

[1] I. Carmesin, K. Kremer, Macromolecules 21 (1988) 2189.
[2] M. Müller, K. Binder, L. Schäfer, Macromolecules 33 (2000) 4568.
[3] D. Chen, W.L. Mattice, Polymer 45 (2004) 3877.
[4] Y. Termonia, Biomacromolecules 5 (2004) 2404.
[5] A. Sokal, in: K. Binder (Ed.), Monte Carlo and Molecular Dynamics Simulations in Polymer Science, Oxford University Press, Oxford, 1955, p. 47.
[6] H. Taketomi, Y. Ueda, N. Gō, Int. J. Pept. Protein Res. 7 (1975) 449.
[7] K.F. Lau, K.A. Dill, Macromolecules 22 (1989) 3986.
[8] G.F. Berriz, E.I. Shakhnovich, Curr. Opin. Colloid Interface Sci. 4 (1999) 72.
[9] Y. Zhang, J. Skolnick, Biophys. J. 87 (2004) 2647.
[10] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, J. Chem. Phys. 21 (1953) 1087.
[11] Y. Duan, P.A. Kollman, Science 282 (1998) 740.
[12] D.L Beveridge, F.M. DiCapua, Annu. Rev. Biophys. Biophys. Chem. 18 (1989) 431.
[13] P.A. Kollman, Chem. Rev. 93 (1993) 2395.
[14] H. Meirovitch, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry, vol. 12, Wiley, New York, 1998, p. 1.
[15] M.N Rosenbluth, A.W. Rosenbluth, J. Chem. Phys. 23 (1955) 356.
[16] F.T. Wall, J.J. Erpenbeck, J. Chem. Phys. 30 (1959) 634.
[17] P. Grassberger, Phys. Rev. E 56 (1997) 3682.
[18] Z. Alexandrowicz, J. Chem. Phys. 51 (1969) 561.
[19] H. Meirovitch, J. Chem. Phys. 89 (1988) 2514.
[20] H. Meirovitch, Chem. Phys. Lett. 45 (1977) 389.
[21] H. Meirovitch, S.C. Koerber, J. Rivier, A.T. Hagler, Biopolymers 34 (1994) 815.
[22] H. Meirovitch, J. Phys. A 16 (1983) 839.
[23] H. Meirovitch, Phys. Rev. A 32 (1985) 3709.
[24] R.P. White, H. Meirovitch, J. Chem. Phys. 119 (2003) 12096.
[25] R.P. White, H. Meirovitch, J. Chem. Phys. 121 (2004) 10889.
[26] S. Cheluvaraja, H. Meirovitch, J. Chem. Phys. 122 (2005) 054903.
[27] H. Meirovitch, Macromolecules 18 (1985) 563.
[28] A.R. Conway, I.G. Enting, A.J. Guttmann, J. Phys. A 26 (1993) L519.
[29] H. Meirovitch, Z. Alexandrowicz, J. Stat. Phys. 15 (1976) 123.
[30] N. Madras, A.D. Sokal, J. Stat. Phys. 47 (1987) 573.
[31] M. Muller, W. Paul, J. Chem. Phys. 100 (1994) 719.
[32] N. Gō, H.A. Scheraga, J. Chem. Phys. 51 (1969) 4751.
[33] M. Karplus, J.N. Kushick, Macromolecules 14 (1981) 325.
[34] H. Meirovitch, J. Chem. Phys. 97 (1992) 5816.