Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 21, Issue Number 6, (2004)

Algorithmic Generation of Freely Jointed Hard Sphere Chains and Properties of their Inertial Tensors

Weizhen Wang[1]*
Y. Bruce Yu[2]**

http://www.jbsdonline.com

Abstract

A statistical algorithm, capable of generating a large number of freely jointed hard sphere chains, is presented. This is the first of a series of algorithms being developed to model unfolded proteins by different modes of hard sphere chains. The aim of these studies is to systematically investigate the effects of different factors, such as atomic radii, bond angles, torsion angles, chain length, etc., on the conformation of unfolded proteins and other random polymers. As continuous models, various types of hard sphere chains enable one to isolate the aforementioned factors one at a time for investigation and thus are advantageous over discrete lattice models. In particular, the freely jointed hard sphere chain model allows one to evaluate the excluded volume effect. As a first step in this endeavor, the average determinant $D(N, r)$ and the average trace $T(N, r)$ of the inertial tensor $\boldsymbol{A}$ of the random chains were calculated at various sphere radii r and chain lengths N. It is found that both the average determinant $D(N, r)$ and the average trace $T(N, r)$ scale linearly with chain length N after logarithmic transformation. However, the critical exponent of $D(N, r)$ increases with r faster than that of $T(N, r)$ as a result of the non-commutativity between the *det* operator and the average operator $< >$. The significance of the algorithm and the results obtained on understanding random polypeptide chains are discussed.

Introduction

Even though the unfolded state comprises half of the protein folding problem, our understanding of it lags far behind that of the folded state (1-3). The reason is obvious. By definition, the unfolded state is made of a large ensemble of strongly fluctuating structures. This feature precludes the description of the unfolded state by a single or a few atomic coordinate sets, as in the case of the folded state. Instead, statistical descriptions should be used for such conformational ensembles. In theoretical studies, short of an analytical treatment, the most direct and thorough approach to this problem is to generate such an ensemble of conformations computationally. Once the ensemble is generated, various aspects of the unfolded state, such as residual structures and solvent accessible surface areas, characters which are singularly important to comprehend protein structure formation, can be investigated. This is the aim of this series of studies.

The abstract construction we adopt to mimic unfolded proteins is the hard sphere chain model in which each unit (an atom or a residue, depending on the resolution) is represented as a hard sphere. This simplicity makes the computation tractable. In the meantime, the hard sphere chain model can incorporate the essential features of a polymer, such as atomic radii, bond angles, torsion angles, chain length and sequence. The hard spheres can also be "softened up" by making the aforementioned parameters adjustable. More importantly, it is a continuous model and hence mimics an unfolded random polypeptide chain more closely than discrete lattice models. Indeed, some very fundamental conclusions about protein structures are

[1]Dept. of Mathematics and Statistics
Wright State University
Dayton, Ohio 45435
[2]Dept. of Pharmaceutics and
Pharmaceutical Chemistry
University of Utah
Salt Lake City, Utah, 84412

*Phone: 937-775-2141
Fax: 937-775-2081
Email: wwang@math.wright.edu

**Phone: 801-581-5133
Fax: 801-585-3614
Email: yby1@utah.edu

based on hard sphere models (4). Our strategy is to build various types of hard sphere chains so that the effect of the aforementioned parameters can be investigated one at a time. In the current work, an algorithm is developed to generate the so-called freely jointed hard sphere chains (FJHSC), the simplest of hard sphere chains. This model incorporates atomic radii and therefore allows one to investigate the excluded volume effect, which is the most important factor in determining the conformation of random polymers (5). Throughout this series of study, at the algorithm development stage for each mode of the hard sphere chain model, the determinant and the trace of the inertial tensor of such a chain will be evaluated. Both quantities are relatively simple parameters of important physical significance, with the determinant related to the rotational entropy of the polymer and the trace to the radius of gyration. At later stages, more complicated characteristics of unfolded proteins, such as residual structure, will be evaluated.

Models and Methods

Description of the FJHSC Model

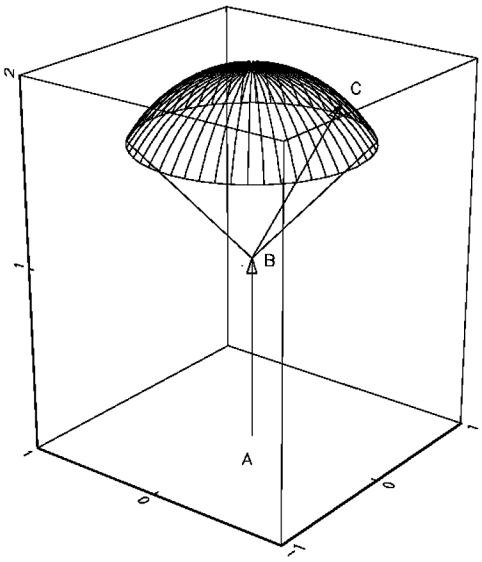

Figure 1: The relationship between two adjacent bonds. For a fixed bond AB, point C, which is one end of bond BC and satisfies distance constraints i) and ii), follows a uniform distribution on the top of the sphere (the "parachute") centered at point B with a radius of 1.

A linear FJHSC is comprised of N hard-sphere units numbered from 1 to N. All the spheres have identical radius, r, which represents the effective van der Waals radius. There are two distance constraints embodying different physicochemical features of a polymer chain (all distances refer to that between the centers of the spheres). i) *The first constraint is that the distance between consecutive units is of fixed length, which is normalized to 1 in this work*. This constraint represents the chemical bond between consecutive units. For a chain of N units, there are N-1 bonds linking the N-1 consecutive unit pairs. ii) *The second constraint is that for any two non-consecutive units, their distance, d, should be no smaller than 2r* (*i.e.*, $d \geq 2r$). This constraint represents the excluded volume effect, i.e., no two units can occupy the same space at the same time. A consequence of the two distance constraints is that the permissible range of r is [0, 1]. Two trivial cases happen when $r = 0$ and 1. When $r = 0$, the FJHSC is reduced to a simple freely jointed chain (mathematically equivalent to a Markov chain), which has been studied in a precursory work (6). When $r = 1$, the chain becomes a straight line and is of no interest here. For r in the range of (0.5, 1), the sum of the radii of two consecutive spheres exceeds the distance between them (i.e., $2r > 1$). This represents the chemical feature that when a covalent bond is formed between two atoms, the bond length is smaller than the sum of the atoms' van der Waals radii. For the FJHSC model, there is no specific restriction on the orientation of the bonds in the three dimensional real space, $\mathbf{R}^3$ (hence the name freely jointed). However, the distance constraint $d \geq 2r$ effectively restricts the rotation of a bond within a cone relative to adjacent bonds (Figure 1).

For a chain made of N hard spheres each of radius r, let (X_i, Y_i, Z_i) be the coordinate of the ith unit with the center of mass of the chain being the origin of the coordinate system. The inertial tensor of the chain is given by the following 3×3 matrix:

$$\boldsymbol{A} = \begin{bmatrix} \sum_{i=1}^{N} m_i (Y_i^2 + Z_i^2) & -\sum_{i=1}^{N} m_i X_i Y_i & -\sum_{i=1}^{N} m_i Z_i X_i \\ -\sum_{i=1}^{N} m_i X_i Y_i & \sum_{i=1}^{N} m_i (Z_i^2 + X_i^2) & -\sum_{i=1}^{N} m_i Y_i Z_i \\ -\sum_{i=1}^{N} m_i Z_i X_i & -\sum_{i=1}^{N} m_i Y_i Z_i & \sum_{i=1}^{N} m_i (X_i^2 + Y_i^2) \end{bmatrix} \quad [1]$$

where m_i is the mass of the ith unit. In this work, $m_i = 1$ for any i. Let $det(\boldsymbol{A})$ and $Tr(\boldsymbol{A})$ denote the determinant and the trace of $\boldsymbol{A}$, respectively. Note that these two quantities vary from chain to chain and thus are random variables. Define

$$D(N, r) = \langle det(\boldsymbol{A}) \rangle \qquad \text{and} \qquad T(N, r) = \langle Tr(\boldsymbol{A}) \rangle \quad [2]$$

where < > means the average value. In this work, r varies from 0 to 0.9 with increment 0.1 while N from 10 to 250 with increment 5.

The Algorithm

At a given value of N and r, the 10,000 conformations were generated using the following statistical simulation. Then $D(N, r)$ and $T(N, r)$ are calculated by averaging the 10,000 values of $det(\mathbf{A})$ and $Tr(\mathbf{A})$, respectively. The simulation, coded in the statistical software *Gauss* (7), was carried out on an Alpha Station, model 200, 4/233 (Digital Equipment Corp.) with a computing speed of 233 megahertz.

1. Without loss of generality, let $(X_1', Y_1', Z_1') = (0, 0, 0)$ and $(X_2', Y_2', Z_2') = (0, 0, 1)$. Suppose that i–1 ($i \geq 3$) points, (X_1', Y_1', Z_1') through $(X_{i-1}', Y_{i-1}', Z_{i-1}')$, have been generated. Now we generate (X_i', Y_i', Z_i').

2. Generate three independent standard normal variables, U, V, W. It is well known that such three random variables have the same distribution in all directions in $\mathbf{R}^3$. Then

$$(U', V', W') = (U, V, W)/(U^2 + V^2 + W^2)^{1/2} \qquad \text{[S-1]}$$

the projection of (U, V, W) onto the unit sphere, is uniformly distributed on the unit sphere.

Remark 1: The standard normal variables can be generated using any commercially available statistical software (e.g., Guass, SAS, Splus, R, etc.) very quickly. Then uniform random variables on the unit sphere are obtained in a straightforward manner by applying the transformation given by Eqn. [S-1]. For example, using the aforementioned computer, one can generate 100,000 points following normal distribution in $\mathbf{R}^3$ and 100,000 points following uniform distribution on a unit sphere in 0.0527 and 0.987 seconds, respectively.

Remark 2: In contrast to the lattice model, our model is a continuous one and the triplet (U', V', W') distributes uniformly on a unit sphere.

3. The coordinate of the ith unit is given by:

$$(X_i', Y_i', Z_i') = (X_{i-1}', Y_{i-1}', Z_{i-1}') + (U', V', W'). \qquad \text{[S-2]}$$

Then distance constraint i) is satisfied. Let $d(i, k)$ be the distance between (X_i', Y_i', Z_i') and (X_k', Y_k', Z_k') for $1 \leq k \leq i - 2$. There are three situations:

3a) If $d(i, k) \geq 2r$ for every $1 \leq k \leq i - 2$ (satisfying distance constraint ii), then (X_i', Y_i', Z_i') is generated by [S-2].

Remark: (X_i', Y_i', Z_i') is obtained by one simulation with the given (X_1', Y_1', Z_1') through $(X_{i-1}', Y_{i-1}', Z_{i-1}')$. One can see that the construction of the ith unit depends on all i-1 previously constructed units. This is different from models based on the Markovian assumption which states that the position of the ith unit depends only on the position of the (i-1)th unit.

3b) If $d(i, k) < 2r$ for any $1 \leq k \leq i - 2$ (violating distance constraint ii), repeat steps 2 and 3 until 3a occurs. Then (X_i', Y_i', Z_i') is generated by [S-2].

Remark: Now (X_i', Y_i', Z_i') is generated by multiple simulations with the given (X_1', Y_1', Z_1') through $(X_{i-1}', Y_{i-1}', Z_{i-1}')$.

3c) If steps 2 and 3 are repeated for many times and step 3a still does not occur, discard the already generated (X_1', Y_1', Z_1') through $(X_{i-1}', Y_{i-1}', Z_{i-1}')$ and restart the procedure from step 1.

Remark: This is an intervention mechanism designed to pull the algorithm out of a dead end. It was implemented through a stopping time t_0, i.e., if (X_1', Y_1', Z_1') through (X_i', Y_i', Z_i') cannot be generated within t_0, then (X_1', Y_1', Z_1') through $(X_{i-1}', Y_{i-1}', Z_{i-1}')$ must be discarded and the process starts over from the first unit. The actual stopping time was chosen to be 12 minutes for the generation of one chain. See more on this in the ***Results*** and ***Discussion*** sections.

4. Suppose that (X_1', Y_1', Z_1') through (X_N', Y_N', Z_N') are generated. The center of mass of the entire random chain, $(\langle X' \rangle, \langle Y' \rangle, \langle Z' \rangle)$, is given by

$$\langle X' \rangle = \frac{1}{N}\sum_{i=1}^{N} X_i', \quad \langle Y' \rangle = \frac{1}{N}\sum_{i=1}^{N} Y_i', \quad \langle Z' \rangle = \frac{1}{N}\sum_{i=1}^{N} Z_i'. \qquad \text{[S-3]}$$

Transform (X_i', Y_i', Z_i') to (X_i, Y_i, Z_i), so that the latter is the coordinate of the ith unit with the center of mass being the origin of the coordinate system, by

$$X_i = X_i' - \langle X' \rangle, \quad Y_i = Y_i' - \langle Y' \rangle, \quad Z_i = Z_i' - \langle Z' \rangle \qquad \text{[S-4]}$$

5. Calculate $det(\boldsymbol{A})$ and $Tr(\boldsymbol{A})$ of (1) with the generated (X_i, Y_i, Z_i)'s.

6. Repeat steps 1-5 10,000 times and calculate $D(N, r)$ and $T(N, r)$ by

$$D(N, r) = \frac{1}{10000}\sum_{j=1}^{10000} det(\boldsymbol{A})_j \,, \quad T(N, r) = \frac{1}{10000}\sum_{j=1}^{10000} Tr(\boldsymbol{A})_j \qquad \text{[S-5]}$$

The algorithm presented here generates flexible, continuous hard sphere chains. Compared with discrete lattice models, a continuous model is more realistic because the conformational space of a random polymer is continuous. Thus a continuous model avoids potential artifacts inherent in lattice models, such as artificial anisotropy. It also allows one to isolate one factor at a time for study. For instance, it is not possible to single out the excluded volume effect in a lattice model.

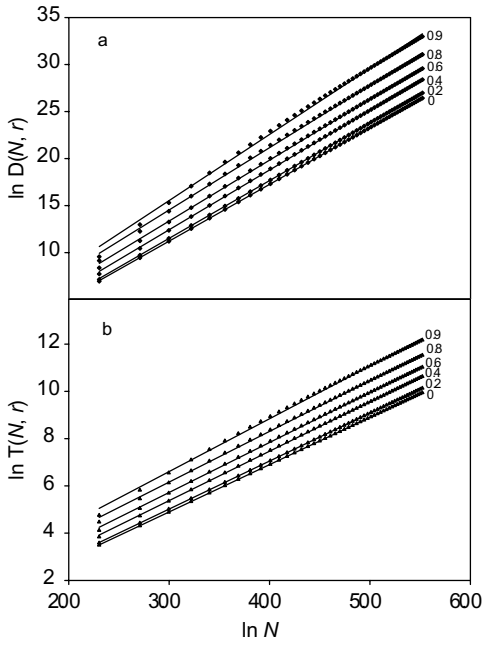

Figure 2: Scaling behavior of the averages of the determinant (**a**) and the trace (**b**) of the inertial tensor of freely jointed hard sphere chains at various sphere radii (r). The dots represent simulation results while solid lines represent linear fits. The slopes of the linear fits give the critical exponents. The value of r is indicated at the upper right corner of each line.

Results

Figures 2a and 2b plot $lnD(N, r)$ and $lnTr(N, r)$ *vs*. lnN, respectively. A linear pattern can be clearly seen for both quantities for large N, i.e.

$$lnD(N, r) = \alpha_1 + \nu_1 lnN \quad \text{and} \quad lnT(N, r) = \alpha_2 + \nu_2 lnN \qquad [3]$$

for large N. Even for small N, the linear equations still fit quite well. Such asymptotic linear relationships mean that both the average determinant and the average trace of the inertial tensor of freely jointed hard sphere chains follow the so-called scaling law (8), with ν_1 and ν_2 being the critical exponents. The values of ν_1 and ν_2 for different r are listed in Table I. From both Figure 2 and Table I, it is clear that the critical exponents for both $D(N, r)$ and $T(N, r)$ increase with r, consistent with the expectation that the average chain conformation would become more expanded as the excluded volume effect becomes more pronounced.

Table I
Scaling constants of $D(N, r)$ and $T(N, r)$ at various sphere radii r.

r	α_1	α_2	ν_1	ν_2
0.9	-5.529	-0.096	7.027	2.235
0.8	-5.322	-0.272	6.623	2.145
0.7	-5.745	-0.485	6.537	2.130
0.65	-5.939	-0.560	6.513	2.124
0.6	-6.049	-0.621	6.473	2.115
0.55	-6.242	-0.710	6.454	2.114
0.5	-6.383	-0.774	6.424	2.107
0.45	-6.492	-0.836	6.385	2.099
0.4	-6.626	-0.890	6.356	2.090
0.35	-6.725	-0.953	6.311	2.080
0.3	-6.785	-0.993	6.255	2.064
0.25	-6.844	-1.034	6.196	2.048
0.2	-6.890	-1.077	6.139	2.032
0.15	-6.838	-1.085	6.078	2.016
0.1	-6.931	-1.122	6.069	2.013
0.05	-6.909	-1.109	6.042	1.998
0	-6.878	-1.111	6.035	2.002

The hard sphere radius also affects the relationship between the two critical exponents. At $r = 0$, $\nu_1 = 3\nu_2$ within simulation error. But as r increases, ν_1 increases faster than ν_2 and therefore $\nu_1 > 3\nu_2$ (Figure 3). This confirms that as a non-linear operator, *det* does not commute with the average operator, $< >$. Otherwise, the following relationship would exist between $<det\boldsymbol{A}>$ and $<Tr\boldsymbol{A}>$:

$$<det\boldsymbol{A}> = det<\boldsymbol{A}> = [\frac{Tr<\boldsymbol{A}>}{3}]^3 = [\frac{<Tr(\boldsymbol{A})>}{3}]^3 \qquad [4]$$

where $<\boldsymbol{A}>$ is the average inertial tensor which, due to the isotropy of freely jointed chains in $\mathbf{R}^3$, is a diagonal 3×3 matrix with all the diagonal terms equal to each other. In the above equation, the second equality holds because of this special feature of $<\boldsymbol{A}>$. The third equality holds because $< >$ and Tr, both linear operators, commute. Therefore, the relationship between $<det\boldsymbol{A}>$ and $<Tr\boldsymbol{A}>$, which leads to $\nu_1 = 3\nu_2$ for any r, rests entirely on the commutativity between $< >$ and *det*. Our simulation results indicate that $< >$ and *det* do not commute, i.e.,

$$<det\boldsymbol{A}> \neq det<\boldsymbol{A}> . \qquad [5]$$

The results on $D(N, r)$ and $T(N, r)$ can be used to calculate the rotational entropy and radius of gyration of FJHSC. The mean square radius of gyration, $<s^2>$, relates to $T(N, r)$ as:

$$<s^2> = T(N, r)/2N . \qquad [6]$$

One of our original motives of this work is to calculate the entropy of overall rotational motions of random polymers, S_{rot}, using ideal gas statistics and compare it with results obtained from experimental measurements. The equation for S_{rot} from ideal gas statistics is (6):

$$S_{rot} = \frac{R}{2}[3 + ln\pi + 3ln\frac{8\pi^2 ml^2 kT}{h^2} + ln<det(\boldsymbol{A})>] \qquad [7]$$

Where m is the mass of each monomeric unit and l is the chemical bond length. R, k, h and T have their usual meanings. For protein molecules, $m \approx 115$Da and $l = 3.63$Å. From our result on $D(N, r)$, the rotational entropy of freely jointed hard sphere chains at 298K is:

$$S_{rot}(N, r) = 16.8R + (\alpha_1 + \nu_1 ln N) R / 2. \qquad [8]$$

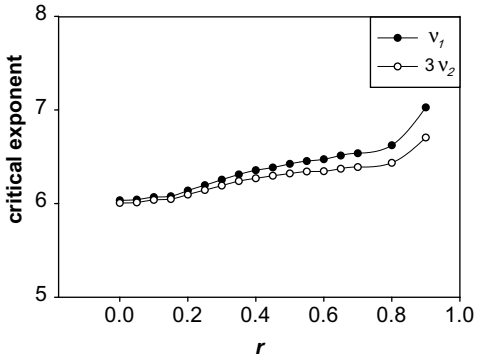

Figure 3: Comparison of the critical exponents of the determinant (ν_1) and the trace (ν_2) at various sphere radii r. Solid symbols represent ν_1 while empty ones represent $3\nu_2$.

From this equation, it can be calculated that for a typical protein molecule of 100 residues (i.e., $N = 100$), the rotational entropy values for $r = 0, 0.5, 0.9$ are $27.3R$, $28.4R$ and $30.3R$, respectively. On the other hand, the measured value of unfolded polypeptide chains in solution is only about one tenth of these calculated values (9). Thus, the results here lend support to the experimental conclusion that ideal gas statistics overestimate the rotational entropy of proteins in aqueous solution (9).

Discussion

To design an algorithm which can generate a large number of freely jointed hard sphere chains is the primary objective of the present work. We tested the algorithm by calculating the properties of the inertial tensor of such chains as functions of chain length. The results indeed demonstrate the expected scaling behavior. A key issue to any statistical algorithm, however, is whether it introduces bias. This question can be answered in a two-step fashion: in the absence ($r = 0$) and in the presence ($r \neq 0$) of excluded volume effect.

In the case $r = 0$, the accuracy of the algorithm can be verified by comparing the simulation results with analytical result on $<s^2>$, the mean square radius of gyration, which is $(N^2 - 1)/6N$ for any $N \geq 3$ (5). Thus,

$$T(N, 0) = (N^2 - 1)/3. \quad [9]$$

Table II
Comparison of simulated and analytical results on $T(N, 0)$ for $3 \leq N \leq 20$.

N	$T(N, 0)$	
	simulated	analytical[a]
3	2.665	8/3
4	4.987	5
5	7.982	8
6	11.65	35/3
7	15.98	16
8	20.99	21
9	26.65	80/3
10	32.97	33
11	39.95	40
12	47.61	143/3
13	55.94	56
14	64.94	65
15	74.60	224/3
16	84.90	85
17	95.87	96
18	107.5	323/3
19	119.8	120
20	132.8	133

[a] analytical result = $(N^2 - 1)/3$.

As can be seen from Table II, excellent agreement is reached between simulated and analytical results on $T(N, 0)$. Thus, in the case $r = 0$, our algorithm samples the conformational space without any bias.

In the case $r \neq 0$, the algorithm presented here is, of course, self-avoiding. Compared with probability-based methods for generating self-avoiding chains (lattice or continuous) (10, 11), the algorithm presented here exhaustively verifies that no steric clash occurs between any non-consecutive pair of hard spheres and hence the avoidance of self-crossing is guaranteed. Probability-based models do not provide such a verification and hence can generated chains with self-crossing, though the probability might be small, depending on the statistical weight factor.

The model presented here is also different from the Rosenbluth self-avoiding on-lattice random walk algorithm (12) and its latter extension to off-lattice flexible random polymers (13, 14). In the Rosenbluth type of algorithm, evaluation of neighboring environment for potential steric clashes takes place at every step during chain growth. Positions for the next move are picked from *permissible* orientations rather than *all* possible orientations. Thus clashes are avoided and no rejections need to be made. However, such algorithm introduces biases which need to be corrected by appropriate weight factors (12-14). In contrast, the algorithm presented here picks positions for each move completely randomly from *all* possible positions. Thus clashes occur and are rejected afterward rather than avoided beforehand. Consequently, no bias is introduced at each growth step and no corrections are needed. The drawback of our algorithm is that sometimes the process can be trapped in a dead end which requires the intervention of a stopping mechanism.

Therefore, if there were any bias, it would come from the selection of the stopping time t_0 in step 3c of the algorithm. The occurrence of step 3c means that the chain is considered to have grown into a dead end and thus the already generated i-1 units, (X_1', Y_1', Z_1') through $(X_{i-1}', Y_{i-1}', Z_{i-1}')$, are discarded. Ideally, one should select $t_0 = +\infty$. This will guarantee that any point on the unit sphere centered at $(X_{i-1}', Y_{i-1}', Z_{i-1}')$ must have distance less than $2r$ with at least one of the i-1 generated points. In other words, the chain has grown into a true dead end and therefore must be abandoned. Of course, such a t_0 cannot be used in practice. For any finite t_0, there is always a possibility that had a longer stopping time been used, a point

on the unit sphere centered at (X_{i-1}´, Y_{i-1}´, Z_{i-1}´) might be found which satisfies all the distance constraints (i.e., the perceived dead end is a fake one). This introduces a bias which favors the more expanded configurations. The question is whether this happens frequent enough to affect the result in a significant way. This can be addressed by varying the stopping time and comparing the results, namely the number of occurrences of step 3c and the coefficients of equation [3]. We carried out such comparisons for $r = 0.9$, the case where step 3c most likely will occur.

a) For 10,000 simulations, step 3c occurs 1013, 881, 927 and 883 times for t_0 = 0.5, 12, 30 and 60 minutes, respectively. Thus, the number of occurrence of step 3c stabilizes at t_0 = 12minutes. This means that for a stopping time of 12 minutes or longer, the intervention mechanism is turned on overwhelmingly by true dead ends.

b) The fitted equations for the relationship between $lnD(N, 0.9)$ and lnN are as follows:

$$\begin{aligned} ln(D(N, 0.9)) &= -5.537 + 7.029lnN, \quad \text{for } t_0 = 0.5 \text{ min.}, \\ ln(D(N, 0.9)) &= -5.530 + 7.027lnN, \quad \text{for } t_0 = 12 \text{ min.}, \\ ln(D(N, 0.9)) &= -5.546 + 7.029lnN, \quad \text{for } t_0 = 30 \text{ min.}, \\ ln(D(N, 0.9)) &= -5.541 + 7.028lnN, \quad \text{for } t_0 = 60 \text{ min..} \end{aligned} \qquad [10]$$

Clearly, the coefficients of the linear fits for various stopping times are essentially the same. Thus, a 120-fold difference in stopping time has no effect on properties of the inertial tensor. Based on these results for $r = 0.9$, we choose 12 minutes as the stopping time for all simulations. With such a stopping time, the bias introduced by the algorithm is negligible.

To demonstrate the scaling behavior, the algorithm presented here is for linear homopolymeric FJHSC in which all the monomers have identical sphere radius, bond length and mass. However, extension of this algorithm to hetero-polymeric FJHSC in which the monomers have differing sphere radii, bond lengths and masses is straightforward. It can also be easily modified to generate branches so that side chains can be incorporated. Thus, various features of unfolded proteins and other random polymers can be added to the algorithm to make the representation more realistic.

Acknowledgement

We thank Prof. Martial Mazars for bringing the Rosenbluth algorithm to our attention. This research was supported by part by an NSF grant to W. W. (grant No. 0308861) and a Faculty Research and Creative Grant to Y. B. Y. from the University of Utah.

References and Footnotes

1. T. P. Creamer, R. Srinivasan and G. D. Rose. *Biochemistry 34*, 16245-16250 (1995).
2. R.V. Pappu, R. Srinivasan and G. D. Rose. *Proc. Natl. Acad. Sci. USA 97*, 12565-12570 (2000).
3. K. A. Dill. *Protein Sci. 8*, 1166-1180 (1999).
4. G. N. Ramachandran and V. Sasisekharan. *Adv. Protein. Chem. 23*, 283-437 (1968).
5. P. J. Flory. *Statistical Mechanics of Chain Molecules, 2nd Ed.* Hanser, New York (1988).
6. Y. B. Yu and W. Wang. *J. Phys. Chem. B 103*, 7676-7680 (1999).
7. Gauss, Version 3.2.31; Aptech Systems, Inc. Maple Valley, WA, (1996).
8. P.-G. de Gennes. *Scaling Concepts in Polymer Physics*. Cornell University Press, Ithaca, NY. (1979).
9. Y. B. Yu, P. L. Privalov and R. S. Hodges. *Biophys. J. 81*, 1632-1642 (2001).
10. D. J. Amit, G. Parisi and L. Peliti. *Phys. Rev. B 27*, 1635-1645 (1981).
11. R. Dickman and C. K. Hall. *J. Chem. Phys. 85*, 4108-4115 (1986).
12. M. N. Rosenbluth and A. W. Rosenbluth. *J. Chem. Phys. 23*, 356-359 (1955).
13. J. I. Siepmann and D. Frenkel. *Mol. Phys. 75*, 59-70 (1992).
14. D. Frenkel, G. C. A. M. Mooij and B. J. Smit. *J. Phys.: Condens. Matter 4*, 3053-3076 (1992).

Date Received: October 16 2003

**Communicated by the Editor
Thomas Cheatham**