

Static and dynamic properties of a new lattice model of polypeptide chains

Andrzej Kolinski, Mariusz Milik, and Jeffrey Skolnick

Citation: *J. Chem. Phys.* **94**, 3978 (1991); doi: 10.1063/1.460675

View online: <http://dx.doi.org/10.1063/1.460675>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v94/i5>

Published by the [American Institute of Physics](#).

Additional information on *J. Chem. Phys.*

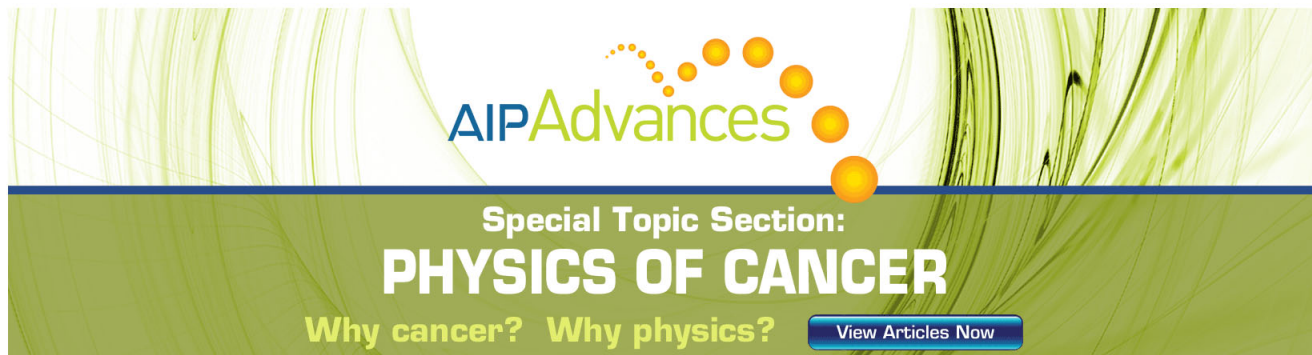
Journal Homepage: <http://jcp.aip.org/>

Journal Information: http://jcp.aip.org/about/about_the_journal

Top downloads: http://jcp.aip.org/features/most_downloaded

Information for Authors: <http://jcp.aip.org/authors>

ADVERTISEMENT



AIPAdvances

Special Topic Section:
PHYSICS OF CANCER

Why cancer? Why physics? [View Articles Now](#)

Static and dynamic properties of a new lattice model of polypeptide chains

Andrzej Kolinski

Department of Chemistry, University of Warsaw 02-093 Warsaw, Poland

Mariusz Miłk and Jeffrey Skolnick^{a)}

Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037

(Received 26 October 1990; accepted 19 November 1990)

The equilibrium and dynamic properties of a new lattice model of proteins are explored in the athermal limit. In this model, consecutive α -carbons of the model polypeptide are connected by vectors of the type $(\pm 2, \pm 1, 0)$. In all cases, the chains have a finite backbone thickness which is close to that present in real proteins. Three different polypeptides are examined: polyglycine, polyalanine, and polyleucine. In the latter two cases, the side chains (whose conformations are extracted from known protein crystal structures) are included. For the equilibrium chain dimensions, with increasing side chain bulkiness, the effective chain length is smaller. The calculations suggest that these model polypeptides are in the same universality class as other polymer models. One surprising result is that although polyalanine and polyleucine have chiral sidechains, they do not induce a corresponding handedness of the main chain. For both polyleucine and polyalanine, the scaling of the self-diffusion constant and the terminal relaxation time are consistent with Rouse dynamics of excluded volume chains. Polyglycine exhibits a slightly stronger chain length dependence for these properties. This results from a finite length effect due to moderately long lived, local self-entanglements arising from the thin effective cross section of the chain backbone.

I. INTRODUCTION

Monte Carlo lattice dynamics have proven to be a very useful tool for investigations of the effect of excluded volume on the static and dynamic properties of polymeric chains,¹⁻³ the statics and dynamics of polymer solutions and melts,^{4,5} and other related problems.⁶

Single polymer chain properties, in particular, the coil-globule transition of the chain, are frequently discussed in the context of polypeptide chain collapse or folding.^{7,8} Recently, we developed a high coordination lattice model of a polypeptide chain which was successfully used in the study of globular protein folding.^{9,10} Since this model is considerably more complex than any of the earlier lattice models of a single polymeric chain, it is reasonable to ask if the general findings of previous lattice chain studies remain valid for the present model. In particular, we will address the following questions: First, what is the effect of side groups on the statics and dynamics of the model chain under high temperature (denaturing) conditions? We modeled these conditions employing the athermal solvent approximation. Consequently, only hard-core repulsive interactions contribute to the mean force of interaction between polymer segments. To some extent, the present model may be viewed as a comb-branching polymer, with appropriate dynamics. The second question concerns the influence, if any, of the built-in chirality of the monomers on the properties of the model. Finally, the effect of varying size side groups is investigated.

For easy reference, we need to consider well defined, as simple as possible, amino acid sequences of these model

chains. Thus, the following three types of polypeptides have been studied: polyglycine, poly-L-alanine, and poly-L-leucine. Polyglycine does not have any side chains, and its α -carbons are achiral. The other two have chiral α -carbons, and their side groups differ substantially in size, which allows us to address the above questions.

Since the present model is generic to the one we are using to simulate globular protein folding,⁹ it is important to find out to what extent its dynamics is physical. If so, the folding pathways we observe can be considered to be more reliable.

This paper is organized as follows. In the next section, we describe the details of the model chain geometry and its physical foundations. The accuracy of the lattice approximation to the polypeptide conformation is analyzed. The Monte Carlo dynamics of the model chains is discussed in Sec. III. Section IV gives a description of the sampling procedure and defines several measured properties of the model polypeptide. In the following two sections (V and VI), we present the results of simulations and discuss the statics and dynamics of these systems. The last section contains concluding remarks.

II. LATTICE MODEL OF POLYPEPTIDE CHAINS

The model described here is a result of a compromise between simplicity and the desire to account for the most important conformational features of polypeptides. The α -carbon representation of the main chain backbone and a spherical representation of the side groups is the option which has been selected. As shown below, the lattice is able to fit with reasonable accuracy the local and global geometry of polypeptides. It is constructed from a subset of simple

^{a)} To whom correspondence should be addressed.

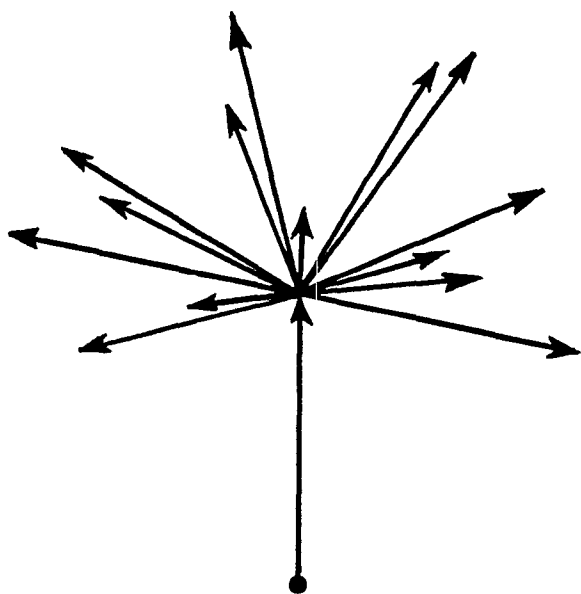


FIG. 1. The set of allowed configurations of two consecutive bond vectors of the backbone of a model 210 lattice polypeptide in the α -carbon representation. There are three (the most expanded) 18-states, one 16-state, four 14-states, two 12-, two 10-, and two 8-states. The corresponding bond angles are: 141.1° , 126.8° , 113.6° , 101.5° , 90° , and 78.5° .

cubic lattice points connected by vectors of the virtual bonds between nearest neighbor α -carbons of the type $(\pm 2, \pm 1, 0)$. For easy reference, it will be called a 210 lattice. The coordination number of the lattice, (the number of possible orientations of the virtual bonds between nearest neighbor α -carbons) is equal to 24. The chain of consecutive α -carbons on this lattice may be visualized as the three-dimensional generalization of a knight move in chess, with some restrictions related to the excluded volume of the backbone and side chains. How accurate is this approximation in comparison to a real protein backbone? The lattice distance (where an integer representation of lattice points is assumed) between two consecutive α -carbons is equal to $5^{1/2}$. If one associates this with the ECEPP¹¹ value for the virtual bond length which is equal to 3.785 \AA , then the local geometry of various structures can be compared. First, let us consider the two most extended (however, with exclusion of the physically impossible collinear one), configurations of a pair of lattice vectors. These "configurational states" may be characterized by the distance $r_{i-1,i+1}^2 = 16$, or 18 in model units. This corresponds to $r_{i-1,i+1} = 6.77$ and 7.18 \AA , respectively, distances which are close to the 6.5 and 7.0 \AA repeat periods for the parallel and antiparallel β sheets.¹² In order to mimic the twist of β structures, the lattice chain has to contain some $r_{i-1,i+1}^2 = 14$ states; thus, one obtains an average which is in quite good agreement with the conformation of real β proteins. A similar level of accuracy holds for α helices, β turns, or loops, as $r_{i-1,i+1}^2$ can assume six allowed values: 8, 10, 12, 14, 16, and 18. We will call these sequences of two virtual α -carbon bonds an 8 state, 10 state, etc. We

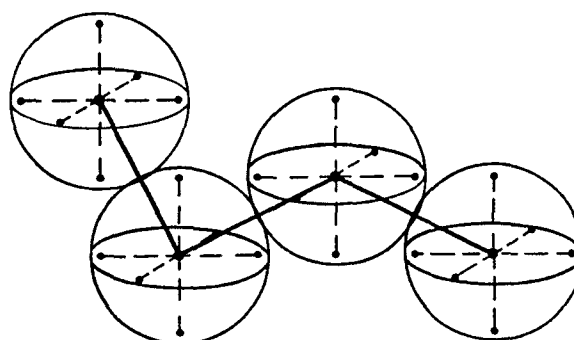
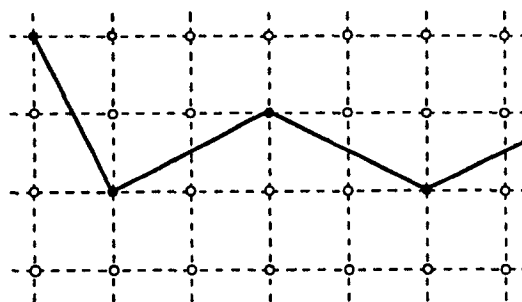


FIG. 2. Schematic representation of the main-chain backbone excluded volume envelope. The distance corresponding to a vector of the type $[1,0,0]$ is about 1.69 \AA .

have *a priori* excluded 2, 4, 6, and 20 states as unphysical; i.e., they do not have their counterparts in real protein chains. In Fig. 1, all possible orientations of two consecutive α -carbon- α -carbon bonds are schematically shown, given a fixed orientation of the first virtual bond.

In order to account for the proper volume of the polypeptide backbone which includes other atoms besides the α -carbons, it is assumed that the six simple cubic (sc) neighbors of the α -carbon 210 lattice site are excluded, an example of which is shown in Fig. 2. This way, the hard core envelope of the main chain is created, where every glycine unit occupies seven points of the underlying sc lattice. This gives a reasonable radius for the excluded volume of the main chain in the range of 1.7 \AA .

The configuration of the side chain of a given amino acid of the polypeptide depends on the local conformation of the main chain. In order to account for the excluded volume of the side chains, we followed the method described recently by Gregoret and Cohen (GC).¹³ Since we need a lattice representation, the method has to be slightly modified and consists of the following steps: First, we exactly reproduced the GC procedure and evaluated the close packing dimensions of side chains for 67 high resolution structures of proteins in the Brookhaven Protein Data Base (PDB).¹⁴ It should be remembered that in GC approach, most of the side chains including those of alanine and leucine, are represent-

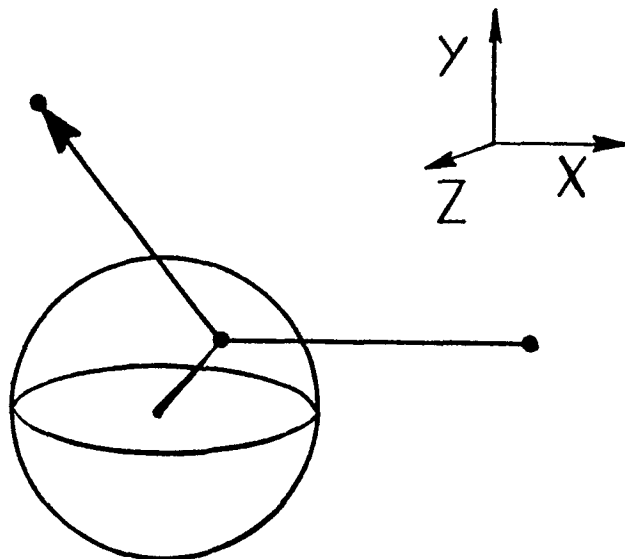


FIG. 3. An example of an alanine side group excluded volume. The distance from the α -carbon to the center of the side group is 1.50 \AA (the corresponding vector for the given example of main-chain conformation, the 16-state, is equal to $([-0.4223, -0.6572, 0.4199])$) and the side group radius is equal to 2.02 \AA , which is in nice agreement with the van der Waals radius of the methylene group.

ed as a single sphere. As a result, we obtained $r_{\text{ALA}} = 2.02 \text{ \AA}$, which is equivalent to the excluded volume radius of the β - CH_3 group in alanine. The average location of the center of mass of the side group for alanine is at a distance R_{ALA} , from the α -carbon and varies over the range from 1.49 to 1.51 \AA . Thus, this distance is almost independent of the main-chain conformation; however, the orientation of side group always depends on the main-chain conformation. The corresponding values for the leucine side group, $-\text{CH}_2\text{CH}(\text{CH}_3)_2$, are $r_{\text{LEU}} = 2.3 \text{ \AA}$ and $R_{\text{LEU}} = 2.36 - 2.48 \text{ \AA}$, depending on the value of $r_{i-1, i+1}^2$. In the second step, we projected every sequence of three amino acids, with the central one being leucine or alanine, onto the nearest possible two bond sequence of the 210 lattice. This defines the side chain orientation and hard-core envelope for every conformation within the lattice approximation. All the sc lattice points inside a given side chain sphere are then considered as excluded to the other side chains as well as to the main chain backbone. Of course, the sphere which mimics the hard-core envelope of the α -carbon under consideration partially overlaps with the sphere of the side group. Both constitute the hard core of a given amino acid. The above representation is quite accurate in spite of the fitting procedure and the averaging of all the proteins over the Brookhaven Protein Data Bank.¹⁴ The proper chirality of all the amino acids are preserved, and presumably, the effect of side chain excluded volume on the conformational spectrum of the model polypeptide is well accounted for. An example of an alanine side group lattice representation is schematically depicted in Fig. 3.

Finally, let us point out that the above lattice model of a

polypeptide (with, of course, a proper representation of side groups of the other amino acids) is able to fit the conformation of all native structures in the PDB with a single set of parameters for a given amino acid. The root-mean-square deviation of α -carbons and centers of mass of the side chains is in the range of 1.5 \AA , which is the level of accuracy of the PDB structures. This means that it is possible to simultaneously reproduce the local as well as global geometry of proteins in the framework of a relatively simple (lattice) model. As a matter of fact, an *a posteriori* test of chirality of the lattice chain representations shows 80% agreement with the corresponding PDB data. Therefore, one may conclude that the model is quite accurate in spite of the assumed lattice representation. (This will be discussed in further detail elsewhere.) Parenthetically, let us note it is implicitly assumed that the side groups conformation (given the backbone conformation) extracted from native structures are equivalent to those in denatured state. Elsewhere, we demonstrate this assumption is in fact very well fulfilled for most of amino acids, and qualitatively is good for all of them.¹⁵

III. MODEL OF DYNAMICS

The high coordination number ($z = 24$) of the 210 lattice allows for a lattice dynamics model which is very close to an off-lattice, bead-and-sticks² model with its essential spike dynamics. Of course, there are substantial lattice restrictions, which are mostly related to the limitation of conformational flexibility in real polypeptides, and therefore they are rather physical.

The dynamics of the entire model chain results from a succession of local conformational transitions. In this sense, it is analogous to the lattice dynamics of a simple cubic (sc) lattice,¹⁶ a diamond lattice,¹⁷ or a face-centered-cubic¹⁸ (fcc) lattice chain. The difference is in the larger number of local jumps involved. Consider a sequence of three beads (α -carbons). For each of the allowed distances between bead 1 and 3, the central bead can be in at least two positions. This condition obtains for conformations $r_{13}^2 = 10, 14, 18$. For $r_{13}^2 = 8$ and 16 , there are four positions of the central bead. For $r_{13}^2 = 12$, there are six possible conformations with fixed beads 1 and 3. Thus, a new α -carbon- α -carbon orientation may be easily created within the chain interior, and the process can be considered to be microscopically reversible, satisfying proper detailed balance. In Fig. 4, the set of allowed conformational transitions is shown for every r_{13} distance. The chain ends have to be treated separately, and for them, the new random orientation of the two end segments is always generated. The model polypeptide chains are constructed as a sequence of $(n + 2)$ beads on the 210 lattice, numbered from 0 to $(n + 1)$. Only n of them can be considered as the α -carbon centers (with side groups for alanine and leucine), while the two end beads serve only to define the conformation of the 1st and the n th amino acid.

The Monte Carlo algorithm works as follows:

(1) The i th bead (one of the $n + 2$ beads) of the model chain is selected at random. For an inner bead, the new conformation is generated by the permutation of two bond vectors (states 10, 14, and 18) or by a random mechanism

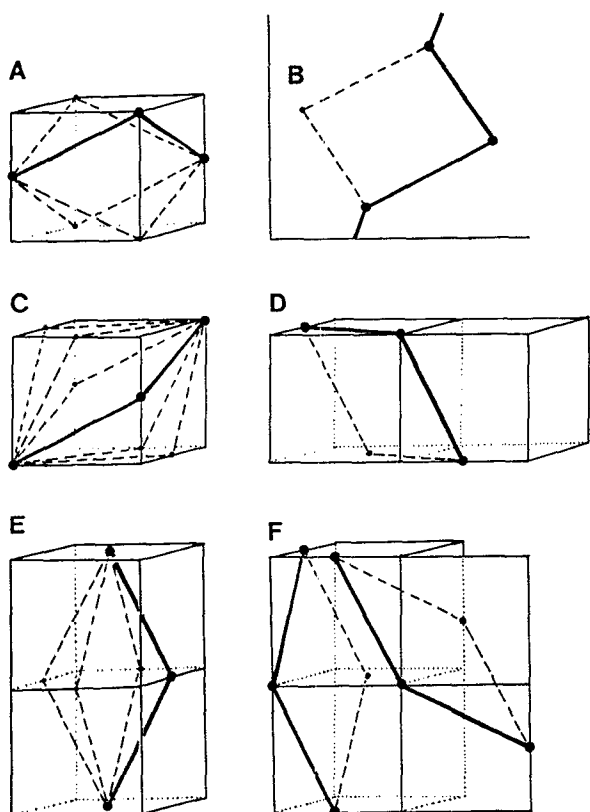


FIG. 4. The set of allowed elementary conformational transitions for the inner segments of model chain, for 8-states (A), 10-states (B), 12-states (C), 14-states (D), 16-states (E), and 18-states (F).

(states 8, 12, and 16). Then, for that obtained trial conformation, it is checked if:

- (a) The conformations of the $(i - 1)$ th and $(i + 1)$ th amino acids belong to the set of allowed conformations.
- (b) The side chains [including the rotated side chain of amino acids $(i - 1)$ and $(i + 1)$] as well as the new position and orientation of the side chain of the i th amino acid do not overlap with themselves or with the rest of the chain.
- (c) The i th α -carbon and its envelope do not overlap with the rest of the chain.

(2) If all the restrictions are satisfied, the new conformation is accepted, and then a new iteration is attempted.

When $i < 1$ or $i > n$, two bond, end rearrangements are invoked with a new conformation generated at random.

A natural unit for the time scale of the model chain dynamics is the time required for $(n + 2)$ attempts at a bead (beads in the case of an end) move. Of course, since the process is controlled by a random number generator, some beads will be not invoked at all in any particular single time unit cycle. The average for every bead over a long time of simulation is, however, the same.

The above procedure requires some comment. It is easy to note that the dynamics is restricted rigorously by the

chains excluded volume. What is less obvious is that topological restrictions are exactly obeyed in that the set of local moves ensures that any portion of the chain cannot pass through another portion of the chain. In other words, one must be careful to avoid bond cutting. The thickness of main-chain hard-core envelope is large enough to prevent such an event for the particular model of local conformational transitions employed here. This way, the model chain dynamics should reproduce the dynamics of a real chain, which is presumably somewhat different than that of an idealized phantom chain (e.g., Rouse¹⁹ dynamics).

IV. SAMPLING PROCEDURE

The initial conformations of non-self-overlapping chains were generated by a separate Monte Carlo algorithm. After that, short equilibration runs were performed. For all three model polypeptides and for each chain length of interest, a few different starting conformations have been prepared. The dynamics of these systems are simulated over a period which was ten (the case of the longest chain, $n = 249$) to several hundred times longer than the longest relaxation time of the chain conformation. The averages from separate independent runs are used to estimate the level of accuracy of the measured properties. The size of the coil is measured by $\langle R_n^2 \rangle$, the mean-square end-to-end distance for the chain of length n , and by $\langle S_n^2 \rangle$, the mean-square radius of gyration of the coil. To facilitate comparison between polyglycine, polyalanine, and polyleucine models only the α -carbons contribute to $\langle S_n^2 \rangle$. Accounting for side chains in the calculation of $\langle S_n^2 \rangle$ changes the obtained values marginally, probably within the statistical error of the simulations. Some higher moments of R_n and S_n were also calculated, which allows for the analysis of the coil shape.

A more detailed description of the local structure of the model chains comes from the inspection of various equilibrium correlation functions. The most important seem to be orientational correlations down the chain. We measured the average cosine of the angle between bond i and j .

$$\langle \cos \Theta_{ij} \rangle = \langle \mathbf{l}_i \cdot \mathbf{l}_j \rangle / l^2 \quad (1)$$

with $l^2 = 5$. The averaging process involves the chain bond index i , time averaging, and averaging over independent runs. The chirality of the main chain backbone has been measured as

$$I_{ij} = \langle (\mathbf{l}_{i-1} \otimes \mathbf{l}_i) \cdot \mathbf{l}_j \rangle; \quad j > 1, \quad (2)$$

where the term in the inner brackets denotes cross product. In the case of a nonchiral chain, the average value of I_{ij} should be equal to zero for all values of i and j . Any nonzero value signals some chirality effect of the side groups on the main-chain conformations.

The dynamics of the model system can be conveniently characterized in terms of various autocorrelation functions. We measured the single bead autocorrelation function $g(t)$, the center of gravity autocorrelation function $g_{c.m.}(t)$, and the autocorrelation function for end-to-end vector $g_R(t)$;

$$g(t) = \langle [\mathbf{r}(t) - \mathbf{r}(0)]^2 \rangle, \quad (3)$$

$$g_{c.m.}(t) = \langle [\mathbf{r}_{c.m.}(t) - \mathbf{r}_{c.m.}(0)]^2 \rangle, \quad (4)$$

$$g_R(t) = \langle [\mathbf{R}_n(t) \cdot \mathbf{R}_n(0)] \rangle / \langle R_n^2 \rangle, \quad (5)$$

with r and $r_{c.m.}$ being the Cartesian coordinate of a single bead (α -carbon) and the center of gravity of the coil, respectively. R_n is the end-to-end vector. The averaging of $g_{c.m.}$ and g_R is performed over time and over multiple runs, while the calculation of g also involves down the chain averaging. The behavior of these autocorrelation functions can be easily compared with theoretical predictions, other simulations and some experimental findings.

V. STATIC PROPERTIES

The average properties of the model polypeptides were obtained by time averaging over sufficiently long trajectories (in most cases, a single trajectory corresponds to the time equivalent of hundreds of τ_R , the terminal relaxation time of the chain conformation) and over a few (2 to 5) separate runs. For two systems, we also performed an additional simulation using a quite different algorithm, in which one side (randomly selected) of the chain is rotated, or bent, by a random mechanism. The equilibrium results coincided (within the statistical uncertainty) with those obtained by dynamic sampling. This may be considered as an additional test of the ergodicity of the Monte Carlo dynamics of these models.

A. Chain dimensions

In Table I, we compare the average characteristics of the model polypeptide coils. The standard deviation of the $\langle S^2 \rangle$ and $\langle R^2 \rangle$ is given for all those cases where at least three independent simulation runs were performed. Inspection of the coil dimensions of various types of polypeptides of equal degree of polymerization shows a substantial increase of $\langle S^2 \rangle$ and $\langle R^2 \rangle$ with increasing molecular mass of the amino acid. Meanwhile, there is an opposite tendency in the changes of $\langle R^4 \rangle / \langle R^2 \rangle^2$. The behavior of this ratio at small n shows that model polyglycine (then polyalanine) more rapidly approaches the long-chain limit of the segment distribution, with the associated proper limiting values of the various

moments. For the ratio of $\langle S^2 \rangle / \langle R^2 \rangle$, the limiting value is close to 0.157 (the value observed in other simulations of lattice chains with excluded volume) in contrast to $1/6$ —the value for an ideal chain. The model polyleucine seems to exhibit the most short chain character, at least for $n = 49$ (via a small value of the $\langle R^4 \rangle / \langle R^2 \rangle^2$ ratio).

Additional confirmation of the above conclusions come from an analysis of the $\langle S_n^2 \rangle$ or $\langle R_n^2 \rangle$ vs n scaling. In Fig. 5, the obtained values of $\langle S_n^2 \rangle$ are compared with the theoretically expected scaling for large n . One may expect an universal behavior of the type $\langle S^2 \rangle \sim n^{2\nu}$. Indeed, the value of 2ν for polyglycine is equal to 1.2, in good agreement with $2\nu = 1.184 \pm 0.04$. This seems to be the limit value for excluded volume, athermal chains. For polyalanine and polyleucine, the exponent over the entire range of n studied is equal to 1.27; however, in the case of polyalanine, the fit to the window $n \gg 99$ is in reasonable agreement with $2\nu = 1.20$. These observations strongly suggest that the model polypeptides belong to the same universality class as other models of "real" (with excluded volume) polymers. However, the large n limit behavior is approached for larger values of n when the size of the side group of the amino acid increases.

B. Bond-bond correlations

In Fig. 6, the logarithm of $\langle \cos(\Theta_{ij}) \rangle$ is plotted vs distance down the chain $|j - i| + 1$. Apparently, the results of simulations can be well fit by straight lines. Therefore,

$$\langle \cos(\Theta_{ij}) \rangle = (|j - i| + 1)^{-a}, \quad (6)$$

with $a = 1.52$ for polyglycine, 1.29 for polyalanine, and 1.18 for polyleucine. In other words, the local correlation of the main chain backbone increases with increasing size of side group.

A somewhat surprising result comes from analysis of the chirality factor for the main chain I_{ij} vs distance down the chain. For all the systems studied, the obtained values of I_{ij} are equal to zero within the statistical noise (some samples are very large), even for short distances down the chain.

TABLE I. Average dimensions of model chains.^a

n	$\langle R_n^2 \rangle$	$\langle S_n^2 \rangle$	$\langle R_n^4 \rangle / \langle R_n^2 \rangle^2$	$\langle S_n^2 \rangle / \langle R_n^2 \rangle$
polyglycine				
24	301.0	45.9	1.39	0.153
49	741.0	114.9	1.46	0.155
99	1749.0	265.5	1.45	0.152
149	2833.0	445.6	1.47	0.157
poly-1-alanine				
49	855.(34.)	128.5(1.3)	1.41	0.150
99	2181.(141.)	330.2(19.2)	1.44	0.151
149	3311.0	516.8	1.50	0.156
199	4794.(444.)	738.7(55.2)	1.50	0.155
249	6415.(878.)	974.0(44.1)	1.48	0.155
poly-1-leucine				
49	963.0	144.2	1.35	0.150
99	2444.(47.)	369.(8.)	1.43	0.151
149	3950.(274.)	620.(22.)	1.47	0.157

^a The numbers in brackets indicates the standard deviation of a single run when three to five independent runs have been performed.

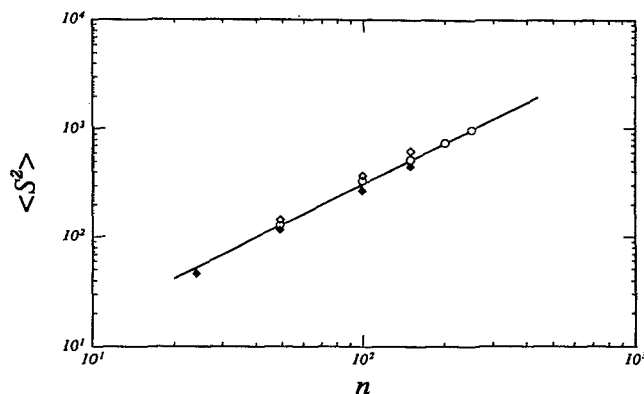


FIG. 5. Log-log plot of $\langle S_n^2 \rangle$ vs chain length n , for polyglycine (solid diamonds), polyalanine (circles), and model polyleucine (diamonds). The solid line corresponds to the expected scaling for a self-avoiding random walk with exponent 1.184.

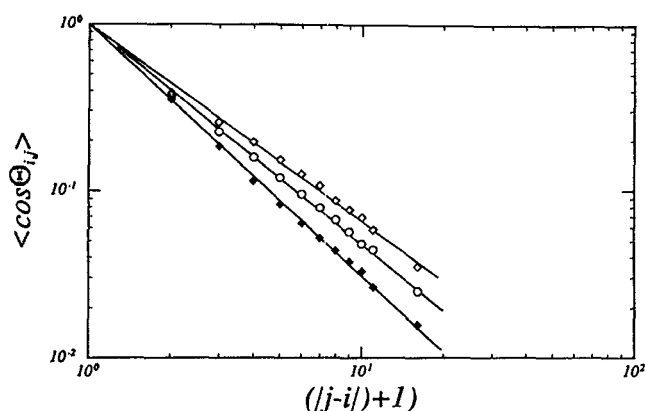


FIG. 6. Semilog plot of $\langle \cos(\Theta_{ij}) \rangle$ vs distance down the chain, for model polyglycine, polyalanine, and polyleucine in the solid diamonds, circles, and diamonds, respectively. Lines correspond to the simple exponential least square fit.

There is no handedness of the main chain induced by the chirality of amino acids under the strongly denaturing conditions described here, i.e., when uniform backbones having chiral side chains are considered. In real polypeptides, the steric clash of the main chain and side chain atoms induces local chirality; such effects are absent in the present model. Therefore, the right-hand twist observed in proteins (α helices, twist of β strands) may be partially induced by attractive interactions (hydrophobic interactions of side groups, hydrogen bonding), or by other tertiary interactions in the native state, or perhaps there must be a high amino acid density (in the native state) in order to enforce the handedness by the steric interactions of side groups. This is not the case in the present simulations.

In Table II, we compare the relative frequency of various conformational states in the three types of model polypeptides. There is a clear increase of population of 18 states (corresponding to expanded β structures) and a slight, but non-negligible, increase of probability of 12 and 14 states (which correspond to wide helical conformations) with increasing size of the side groups. This is accompanied by a decrease in population of the 8 states and to a lesser extent of

TABLE II. Probability of various conformational states in model chains.^a

Chain type	$\langle R_n^2 \rangle$	10	12	14	16	18
Polyglycine	0.1259	0.1357	0.1380	0.2913	0.0760	0.2331
Poly-l-alanine	0.1112	0.1275	0.1358	0.3009	0.0796	0.2450
Poly-l-leucine	0.0944	0.1222	0.1404	0.3037	0.0791	0.2602

^a All data for $n = 149$. The statistical uncertainty, as well as the differences between the figures for various values of n , are conservatively estimated to be below 1%.

10 states. This way, excluded volume, even in very expanded conformation, enforces a bias towards more expanded, β sheetlike global configurations.

VI. DYNAMIC PROPERTIES

The dynamics of the single polymeric chain, when hydrodynamic interactions are ignored (which is also the case of present model) should be similar to dynamics of a bead-and-spring Rouse¹⁹ model. The Rouse model is solvable analytically and predicts that the terminal relaxation time, (which is equal to relaxation time for end-to-end vector of the chain) to be proportional to n^2 , and the diffusion coefficient should be proportional to n^{-1} . Marginally, let us note that it has been proven that freely jointed²⁰ chains which are very similar to high coordination number lattice chains and lattice chain dynamics,²¹ without excluded volume and topological restrictions, is homeomorphic to Rouse dynamics as far as the time scales considered are large in comparison to the single bond relaxation time. The introduction of excluded volume and topological restrictions is essentially possible only in simulated models and seems to substantially increase the terminal relaxation time. Most of the simulations on simple lattice models suggest that for this kind of model, the terminal relaxation time should scale at least like $n^{2.2}$, with 2.2 being the limiting (large n) value of the exponent.^{18,22,23} Approximate theoretical estimations²³ of this exponent agree with the simulations. If one uses the original Rouse model, $\tau_R \propto n \langle S^2 \rangle$ and thus the dynamic scaling hypothesis with $\langle S^2 \rangle \sim n^{1.2}$ gives $\tau_R \propto n^{2.2}$.

First, we checked if the present models exhibit a well defined terminal relaxation time. If so, one should observe, after a very short initial relaxation period, an exponential decay of $g_R(t)$;

$$g_R(t) \sim \exp(-t/\tau_R) \quad (7)$$

with τ_R the longest relaxation time of the chain conformation. Of course, there are additional internal modes that contribute at short time. Figure 7 shows that there is indeed a very well defined longest relaxation time and that the other modes relax very fast. Therefore, we may use a simplified method to estimate τ_R ; namely, the time when $g_R(t)$ decays to $1/e$ of its initial value. As a matter of fact, this estimate provides the same degree of accuracy as values obtained from slopes of semilog plots of $g_R(t)$ vs t . The log-log plot of τ_R vs chain length n is given in Fig. 8.

In Fig. 9, we plot examples of $g(t)$ and $g_{c.m.}(t)$ autocorrelation functions. Again, it is qualitatively the picture expected for Rouse-like dynamics. The diffusion coefficient D , which is defined as

$$D = \lim_{t \rightarrow \infty} (g_{c.m.}(t)/6t) \quad (8)$$

may be extracted from log-log plots of $g_{c.m.}(t)$ against t , from $g_{c.m.}(t)$ vs t plots (slope), or just from the ratio $g_{c.m.}(t)/6t$ at large times, i.e., larger than the terminal relaxation time. A consistent method we use employs the ratio at the times when $g_{c.m.}(t) = 2\langle S_n^2 \rangle$. The obtained values of D are plotted in Fig. 10 vs chain length on a log-log scale.

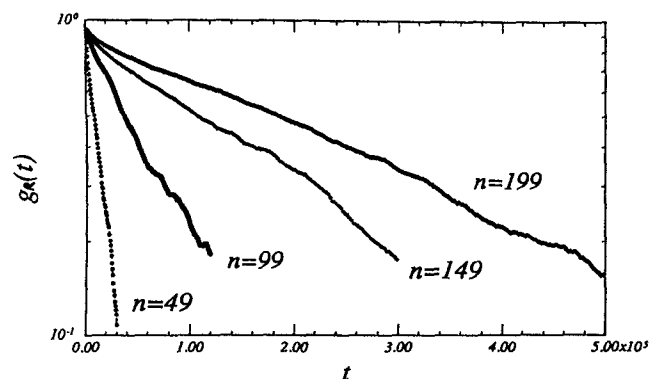


FIG. 7. Semilog plots of $g_R(t)$ vs time for model polyalanine of various chain lengths. The longest relaxation time was estimated as that when $g_R(t)$ decreases to $1/e$ of its initial value. The average slopes of the curves in this figure reflect the longest relaxation time as well.

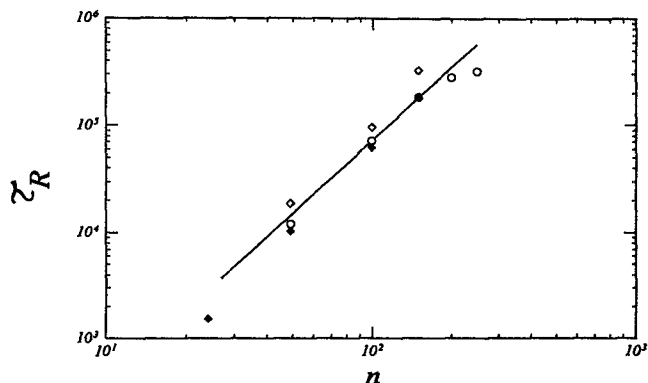


FIG. 8. Log-log plot of the longest relaxation time τ_R vs chain length n , for model polyglycine (solid diamonds), polyalanine (circles), and poly-leucine (diamonds). The solid line with a slope of 2.2 is drawn for comparison with the expected scaling.

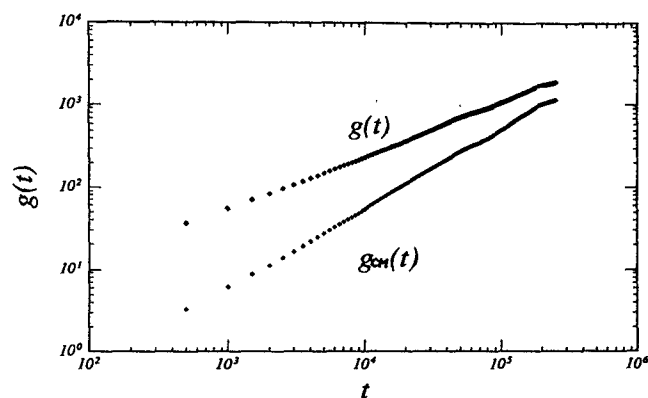


FIG. 9. Log-log plot of a single α -carbon autocorrelation function $g(t)$ and center of mass autocorrelation function $g_{c.m.}(t)$ vs time for model polyalanine having a chain length of $n = 99$.

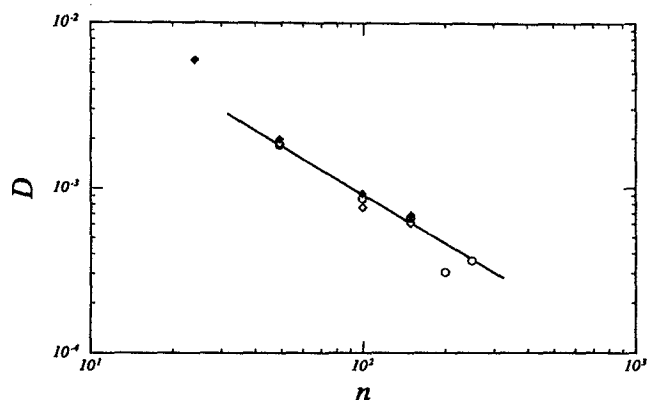


FIG. 10. Log-log plot of the diffusion coefficient D vs time for model polyglycine (solid diamonds), polyalanine (circles), and poly-leucine (diamonds). The solid line corresponds to $D \sim 1/n$ scaling.

Prior to the discussion of chain length and amino acid-type dependence of these global dynamic properties, we should note that side groups have some effect on local dynamics. In Table III, we compare the acceptance ratio of local conformational jumps. With an increase of side chain size, there is a considerable decrease of acceptance ratio, and the dynamics is slowed down due to local conformational stiffness. Part of the observed increase of τ_R (and decrease of D) with increasing size of the side group seen in Fig. 8 (Fig. 10) could be connected to the decreasing frequency of elemental jumps. The lines drawn in Fig. 8 and in Fig. 10 have slopes 2.2 and -1 , respectively. Therefore, the dynamics of model nonglycine, polypeptides seems to be consistent with the dynamics of simpler lattice models of single polymer chains. The accuracy of our data does not allow for an exact estimation of these exponents. For the case of polyglycine chains, there seems to be a somewhat stronger chain length dependence. The fit of τ_R exhibits the exponent 2.45 ± 0.1 , and for D , the exponent seems to be -1.1 ± 0.2 . We may rationalize this by a finite length effect, which is enhanced for polyglycine by local, moderately long lived self-entanglements of the chain backbone. Thin chains having just a single point backbone (with appropriate methods of accounting for topological requirement of non-self-intersection) exhibit this effect even more dramatically—segments get caught much like hooks, therefore enforcing long lived contacts.

TABLE III. Fractions of accepted moves for model chains.^a

Chain type	One bead jumps	End flips
Polyglycine	0.4388	0.9128
Poly-L-alanine	0.4341	0.8720
Poly-L-leucine	0.3940	0.8164

^aData for all values of n exhibit a statistical uncertainty well below 1%.

These entanglements are more frequent and of a longer lifetime for polyglycine than for two remaining models due to the small backbone thickness.²⁴ The effect can artificially enhance the chain length dependence of the relaxation time in the range of relatively short chains.

VII. CONCLUSION

In this work, we described a discretized model of polypeptide chains. The approach is based on an appropriate high coordination lattice representation of the main chain backbone and the side chain excluded volume envelope.

It has been shown that side groups contribute to the excluded volume of the entire polymer, increasing its coil size. However, these model chains belong to the same universality class as other simpler lattice chains with excluded volume. We have shown that in the absence of an attractive interaction between side groups (and/or when coil is expanded) there is no handedness of the main chain backbone induced by the built-in chirality of model aminoacids.

The dynamic properties of model polypeptides are also consistent with the dynamics of simple lattice chains. The results of our Monte Carlo lattice dynamic simulations show that the dynamics is essentially similar to that of a Rouse chain, with some enhancement of the chain length dependence of the terminal relaxation time. The results are consistent with $\tau_R \sim n^{2.2}$ and $D \sim n^{-1}$, in qualitative agreement with other lattice dynamic studies of single polymeric chains.

Consequently, we may expect that the lattice dynamics of similar 210 lattice models for the globular proteins (of course, with a manifold of side chains sizes in accord with the amino acid sequence) can be considered as a reasonable approximation to the dynamics of real proteins. The time scale when various short and long range factors are introduced will probably be somewhat distorted, but Rouse-like dynamics should be preserved.

ACKNOWLEDGMENT

This research was supported in part by a grant from the Polymer Program of the National Science Foundation.

- ¹ A. Baumgartner, *Ann. Rev. Phys. Chem.* **35**, 419 (1984).
- ² A. Baumgartner, in *Application of the Monte Carlo Method in Statistical Physics* (Springer, Heidelberg, 1984).
- ³ A. Kolinski, J. Skolnick, and Robert Yaris, *J. Chem. Phys.* **85**, 3585 (1986).
- ⁴ J. Skolnick and A. Kolinski, *Adv. Chem. Phys.* **78**, 223 (1990), and references therein.
- ⁵ T. Pakula and S. Geyler, *Macromolecules* **20**, 2909 (1987); **21**, 1670 (1987).
- ⁶ *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer, Berlin, 1986).
- ⁷ O. B. Ptitsyn and V. A. Finkelstein, *Q. Rev. Biophys.* **13**, 339 (1980).
- ⁸ J. Skolnick and A. Kolinski, *Ann. Rev. Phys. Chem.* **40**, 207 (1989).
- ⁹ J. Skolnick and A. Kolinski, *Science* **250**, 1121 (1990).
- ¹⁰ J. Skolnick and A. Kolinski, *J. Mol. Biol.* (submitted for publication).
- ¹¹ G. N. Ramachandran and V. Sasisekharan, *Adv. Protein Chem.* **23**, 283 (1968).
- ¹² C. Ghelis and J. Yon, *Protein Folding* (Academic, New York, 1982).
- ¹³ L. M. Gregoret and F. E. Cohen, *J. Mol. Biol.* **211**, 959 (1990).
- ¹⁴ F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, *J. Mol. Biol.* **112**, 535 (1977).
- ¹⁵ L. Pielka, M. Milik, A. Kolinski, and J. Skolnick (manuscript in preparation).
- ¹⁶ A. Kolinski, J. Skolnick, and R. Yaris, *J. Chem. Phys.* **86**, 7164, 7174 (1987), and references therein.
- ¹⁷ F. Geny and L. Monnerie, *J. Poly. Sci. Polym. Phys. Ed.* **17**, 131, 147 (1979).
- ¹⁸ C. Stokely, C. C. Crabb, and J. Kovac, *Macromolecules* **19**, 860 (1986).
- ¹⁹ P. E. Rouse, *J. Chem. Phys.* **21**, 1272 (1953).
- ²⁰ R. A. Orwoll and W. H. Stockmayer, *Adv. Chem. Phys.* **15**, 305 (1969).
- ²¹ K. Iwata and M. Kurata, *J. Chem. Phys.* **50**, 4008 (1969).
- ²² A. Kolinski, J. Skolnick, and R. Yaris, *J. Chem. Phys.* **86**, 1567 (1987).
- ²³ P. G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell, Ithaca, 1979).
- ²⁴ A. Sikorski, A. Kolinski, and J. Skolnick (unpublished).