Folding in two-dimensional off-lattice models of proteins

Mai Suan Li and Marek Cieplak

Institute of Physics, Polish Academy of Sciences, Allja Lotnikow 32/46, 02-668 Warsaw, Poland

(Received 2 March 1998; revised manuscript received 2 June 1998)

Off-lattice proteinlike models are constructed in two dimensions so that their native states are close to an on-lattice target. The Hamiltonian involves the Lennard-Jones and harmonic interactions. The native states of these sequences are determined with a high degree of certainty through Monte Carlo processes. The sequences are characterized thermodynamically and kinetically. It is shown that the rank-ordering-based scheme of the assignment of contact energies typically fails in off-lattice models even though it generates high stability of on-lattice sequences. Similar to the on-lattice case, Go-like modeling, in which the interaction potentials are restricted to the native contacts in a target shape, gives rise to good folding properties. Involving other contacts deteriorates these properties. [S1063-651X(99)05301-5]

PACS number(s): 87.15.By, 87.10.+e, 71.28.+d, 71.27.+a

I. INTRODUCTION

Understanding of the folding process of proteins is one of the current challenges in molecular biophysics. Many insights into the nature of folding have been provided by studying lattice models in which a protein is represented by a chain of beads on a hupercubic lattice (see, for instance, Ref. [1]). A more realistic modeling of proteins, however, requires considering off-lattice systems. Simple off-lattice systems have been discussed recently by Irback *et al.* [2] and Klimov and Thirumalai [3]. The former authors have studied a model with two kinds of residues and they have found that very few such sequences give rise to a rapid folding.

In this paper we focus on ways to design off-lattice sequences that form good folders. Specifically, we consider a two-dimensional (2D) target shape and assign two models of interaction energies between the beads. The first model is a generalization of the Go-like approach [4] to off-lattice situations. The second model, on the other hand, is a generalization of the rank-ordering-based assignment of the couplings [5,6]. When the beads are constrained to be located on sites of a lattice, both models provide sequences which are good folders. Here, we demonstrate that this may not be true for off-lattice models. Namely, an extended character of the interactions, that is necessarily present in off-lattice Hamiltonians, gives rise to differing levels of frustration in the two models and leaves only the Go-like sequences as good folders. This is compatible with the principle of minimum frustration proposed by Bryngelson and Wolynes [7].

We formulate our models in the context of the Hamiltonian used by Iori, Marinari, and Parisi [8] (IMP) in which monomers interact via the Lennard-Jones potential. The amplitudes of the attractive part between beads i and j of this potential, A_{ij}, representing a residue-dependent interaction, are quenched random variables. Additionally, the monomers are tethered sequentially along the chain by means of harmonic interactions. IMP and Struglia [9] have demonstrated the existence of a compact phase in a 3D version of the model. In these studies the dynamics have been defined in terms of a Monte Carlo process. The true ground state—the native state—however, typically is not known (as in IMP) or there is a substantial degree of uncertainty whether the state

assumed to be native is indeed the ground state (as in the paper by Struglia [9]).

Studies of the dynamics of folding require knowing the precise shape of the native state conformation without which the folding time cannot be defined. For small scale lattice models, the ground state may be obtained by an exact enumeration of conformations. This method, however, does not apply to off-lattice models. Here we present 2D Lennard-Jones sequences of 16 monomers in which the ground state is known accurately and with a high degree of certainty. We then provide some basic characterization of these sequences, obtained through the Monte Carlo procedure.

The construction of the model sequences is presented in Sec. II. We consider the lattice target shown in Fig. 1(a). In order to generate a Go-like sequence G, we assign the A_{ij} of the Lennard-Jones potential to be uniform in the native contacts and zero in the non-native contacts, i.e., all pairs of monomers that do not form contacts in the native state interact only through a short-range repulsion. The second model sequence, R', is constructed by generating Gaussian A_{ij} and adopting the rank-ordering scheme introduced recently [5,6] in the context of lattice models. Sequence R' is an off-lattice analog of the sequence R discussed in Refs. [5,6]. The principle here is to allocate the most strongly attractive A_{ij} to those pairs of monomers which form contacts in the target compact lattice conformation. The true ground state of the sequence will necessarily be off lattice but its shape will be close to the lattice target—it can be viewed as a somewhat distorted lattice target conformation.
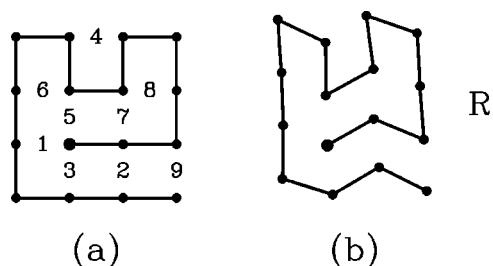

FIG. 1. (a) Assigment of nine strongest couplings A_{ij} to native contacts for sequence R'. The numbers indicate the relative strengths of the contacts. (b) The native conformation of sequence R'.

In Sec. III we describe the procedure adopted to determine the ground state. Basically, we form self-avoiding walks (SAW) in continuum space around all of the 69 maximally compact lattice conformations as the starting configuration of the sequence. We shall denote such starting SAW as CSAW to emphasize closeness to compact structures. We then perform Monte Carlo quenches to reach local energy minima. The rank-ordered allocation of the couplings ensures that the lowest energy minima will be compact. For sequence R' we have shown that the lowest energy state is obtained from CSAW's which are near the target structure. We demonstrate that the use of arbitrary SAW's, i.e., which are not CSAW's, leads to local minima of higher energies. For sequence G both the SAW's and CSAW's starting configurations easily lead to the native state which has the target compact shape. Our conclusion is that it is easier to find the native state for a good folder (G) compared to a bad folder (R'). We have found also that only sufficiently large values of the spring constant k may guarantee the self-avoidance properties of the chain.

We characterize geometries of conformations in terms of a certain distance away from a reference conformation. In Sec. IV we determine the probability to stay in the basin and the folding temperature. Both quantities are calculated by adopting a well defined characteristic basin size as obtained by the shape distortion approach [10]. In Sec. V we present results on the specific heat and structural susceptibility for R' and G and demonstrate that G is a good folder whereas R' is a bad folder. The bad folding properties of R' are in sharp contrast to what was found for the lattice sequence R [5,6] and are due to the presence of many long-ranged couplings which impose conflicting constraints. Sequence G is a good folder because there are many fewer constraints to satisfy.

II. MODELS OF INTERACTIONS

Following IMP [8], we consider a self-interacting heteropolymer in 2D described by the Hamiltonian given by

$$H = \sum_{i \neq j} \left\{ k(d_{i,j} - d_0)^2 \delta_{i,j+1} + 4 \left[\frac{C}{d_{i,j}^{12}} - \frac{A_{ij}}{d_{i,j}^6} \right] \right\}, \quad (1)$$

where i and j range from 1 to the number of beads, N, which in our model is equal to 16. The distance between the beads, $d_{i,j}$, is defined as $|\vec{r}_i - \vec{r}_j|$, where $\vec{r}_i$ denotes the position of bead i. The harmonic term in the Hamiltonian with the spring constant k couples the adjacent beads along the chain. The remaining terms represent the Lennard-Jones potential. In [8] A_{ij} is chosen as $A_{ij} = A_0 + \sqrt{\epsilon} \eta_{ij}$, where A_0 is constant and η_{ij}'s are Gaussian variables with zero mean and unit variance; ϵ controls the strength of the quenched disorder. The case of $\eta_{ij} = 0$ and $A_0 = C$ would correspond to a homopolymer with the standard Lennard-Jones interaction used in simulations of liquids. We adopt the units in which $C = 1$ and consider k to be either equal to 1 or to 25. We have found that the first choice, which has been used by IMP [8], may lead to local energy minima in which the polymer is self-intersecting.

For $N = 16$, there are 120 Lennard-Jones possible couplings between the monomers. The basic choice for the values of A_{ij} that we shall use here is shown in Table I. Here, all of the A_{ij}'s are positive, which corresponds to attraction. Table I also indicates a rank of a given attractive coupling if these are rank ordered from the strongest to the weakest attraction. In nine of the couplings, the attraction is enhanced by making the corresponding A_{ij} bigger than one. These were chosen to coincide with the contacts present in the lattice target R' shown in Fig. 1(a). The strongest attraction was assigned to be between beads 1 and 12 and the relative strengths of other native attractions are indicated in the figure. The remaining 111 couplings are assigned A_{ij} with values which are smaller than 1. Overall the mean value A_0 is equal to 0.784 and the dispersion to 0.205.

The parameter d_0 corresponding to the equilibrium distance in the harmonic potential is chosen to be equal to 1.16, which is close to the equilibrium position of the average Lennard-Jones potential, $(2C/A_0)^{1/6}$. The target lattice shapes are built on a lattice with this lattice constant. For the Go-like sequence (G) we set $A_{ij} = 1$ for native contact and 0 for non-native ones. For this sequence we choose $d_0 = 2^{1/6}$. The qualitative results do not depend, however, on the choice of d_0.

III. LOCAL MINIMA AND NATIVE STATES

In order to find spectra of the local energy minima we use the Monte Carlo procedure with local updating moves. Monomers are moved randomly within a circle (the radius of circle varies from $0.0025d_0$ to $0.01d_0$) away from their previous positions. Random quenching by starting from a SAW or a CSAW typically yields a local energy minimum within of order 2^{18} steps. In the minimum, the force acting on any of the monomers is at most of order 10^{-7} in units of the characteristic coupling (of order 1) divided by $(2C/A_0)^{1/6}$. Simulated annealing runs of comparable number of steps yielded similar results. The results reported on in this paper are based on the quenching procedure.

The starting conformations were obtained by placing monomers within a circle of radius $0.3d_0$ away from a SAW generated on the lattice or away from a maximally compact conformation on the lattice CSAW. For each of the 69 maximally conformations, 50 CSAW's were generated and the results were compared with those obtained based on 500 SAW's.

We observe that there is a substantial gap, of order 4, between the minima obtained from CSAW's around the target and those obtained from all other starting configurations.

As to the choice of the elastic constant k, we have found that $k = 1$ is not very physical because the distances between consecutive monomers are not kept sufficiently rigid. Furthermore, we have observed that in states which are not local energy minima, the polymer conformations may become self-intersecting. This phenomenon was also observed in Ref. [11] for shorter chains. Thus a stronger k is needed and we focus on $k = 25$.

The energy histograms for the R' and G sequences for $k = 25$ are shown in Figs. 2 and 3, respectively. In the case of R', the SAW configurations happen to yield results comparable to the CSAW configurations, but the statistical frequency of success in finding a low energy state favors the CSAW-based approach. The maximal distance between

TABLE I. Values of the A_{ij} for sequence R'. The rank of a coupling is indicated to the right of its value.

i	j	A_{ij}	Rank	i	j	A_{ij}	Rank	i	j	A_{ij}	Rank	i	j	A_{ij}	Rank
1	2	0.946554	19	3	5	0.581629	105	5	12	0.667698	79	8	15	0.758490	57
1	3	0.858677	37	3	6	0.598687	101	5	13	0.740622	62	8	16	0.718888	69
1	4	0.810776	47	3	7	0.513536	118	5	14	0.661444	83	9	10	0.913366	28
1	5	0.949365	17	3	8	0.778139	50	5	15	0.861819	36	9	11	0.741376	61
1	6	0.514208	117	3	9	0.746676	59	5	16	0.744776	60	9	12	0.664109	82
1	7	0.644279	91	3	10	0.666580	80	6	7	0.767029	54	9	13	0.968419	11
1	8	1.337928	5	3	11	0.835338	41	6	8	0.638391	93	9	14	0.822724	45
1	9	0.725085	65	3	12	0.722943	67	6	9	1.375767	4	9	15	0.969623	10
1	10	0.802427	48	3	13	0.518711	116	6	10	0.676479	76	9	16	0.850939	39
1	11	0.763514	55	3	14	0.660732	85	6	11	0.567934	109	10	11	0.935812	22
1	12	1.498411	1	3	15	0.852925	38	6	12	0.571524	108	10	12	0.951547	16
1	13	0.899355	30	3	16	1.050544	9	6	13	0.658820	86	10	13	0.521468	114
1	14	1.396376	3	4	5	0.878341	33	6	14	0.512926	119	10	14	0.930277	25
1	15	0.723510	66	4	6	0.763190	56	6	15	0.612393	99	10	15	0.889378	31
1	16	0.655780	87	4	7	1.269149	8	6	16	0.628334	95	10	16	0.590144	104
2	3	0.720592	68	4	8	0.577229	106	7	8	0.670401	78	11	12	0.932459	23
2	4	0.541154	112	4	9	0.614099	96	7	9	0.647826	89	11	13	0.528390	113
2	5	0.770339	53	4	10	0.918839	27	7	10	0.642560	92	11	14	0.921718	26
2	6	0.938498	21	4	11	0.597736	102	7	11	0.965115	12	11	15	0.770826	52
2	7	1.291634	7	4	12	0.571611	107	7	12	0.834125	42	11	16	0.955652	14
2	8	0.696944	72	4	13	0.561490	110	7	13	0.753538	58	12	13	0.629158	94
2	9	0.603317	100	4	14	0.725428	64	7	14	0.561397	111	12	14	0.884585	32
2	10	0.832387	43	4	15	0.954825	15	7	15	0.712665	70	12	15	0.949139	18
2	11	0.672241	77	4	16	0.511493	120	7	16	0.864981	35	12	16	0.932023	24
2	12	0.821738	46	5	6	0.687186	74	8	9	0.943903	20	13	14	0.848886	40
2	13	0.644946	90	5	7	0.665046	81	8	10	0.660797	84	13	15	0.613872	97
2	14	0.689698	73	5	8	0.686304	75	8	11	1.301709	6	13	16	0.734204	63
2	15	1.409250	2	5	9	0.771102	51	8	12	0.903380	29	14	15	0.519439	115
2	16	0.655129	88	5	10	0.595747	103	8	13	0.612414	98	14	16	0.963250	13
3	4	0.781090	49	5	11	0.707574	71	8	14	0.825123	44	15	16	0.866528	34

monomers does not exceed d_0 by more than 4% in any local energy minimum, which demonstrates very good self-avoiding properties.

In the case of sequence G, about 25% of CSAW's trajectories and 12% of SAW's trajectories lead to the native conformation. Thus it is easier to find the native state for G than for R'.

The native state conformation of sequence R' is shown in Fig. 1(b). The x-y coordinates of the beads of this conformation are listed in Table II. They are used in the studies of kinetics and thermodynamics properties. The native state of sequence G has exactly the lattice shape as shown in Fig. 1(a) because it is only the native contact energies here that one needs to minimize.

As we shall show later, several tasks, like the determination of the folding temperature, are facilitated by introducing a notion of a geometrical distance between two conformations. A convenient definition of the distance δ_{ab} between two conformations a and b is provided by

$$\delta_{ab}^2 = \min \frac{1}{N} \sum_{i=1}^{N} |\vec{r}_i^a - \vec{r}_i^b|^2, \qquad (2)$$

where $\vec{r}_i^{a,b}$ denotes the position of monomer i in conformation $a(b)$. The minimization is performed over translations,
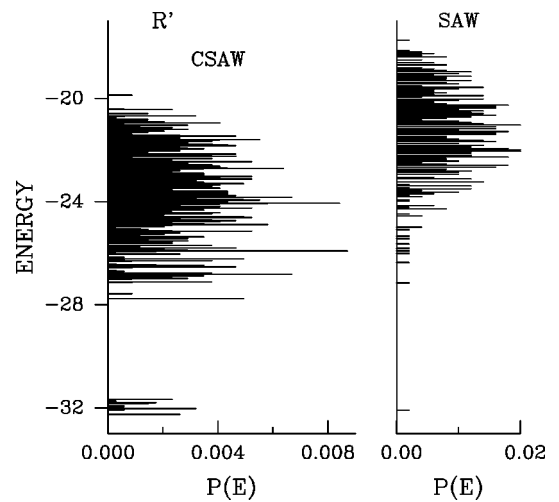

FIG. 2. Density of local minima for sequence R' and $k=25$. The left part corresponds to those obtained from deformed compact cells as starting conformations, whereas the right part corresponds to the case when starting conformations are arbitrary self-avoiding chains. The energy gap is equal to 0.15. (For $k=1$ above $E \approx -23.93$ the distance between the beads may become bigger than the equilibrium distance of the Lennard-Jones potential by 10%. This happens in about 1.4% of all conformations.)
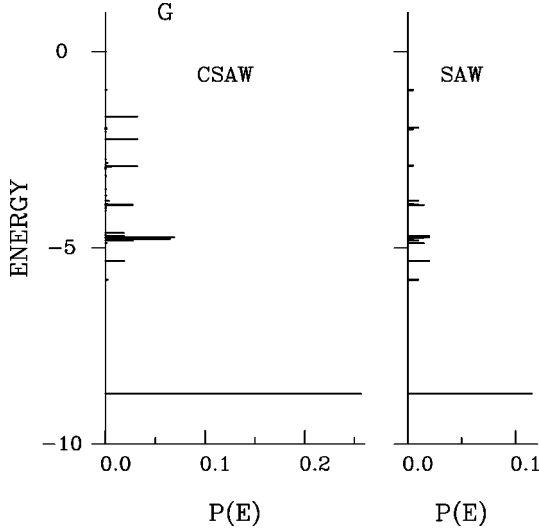
FIG. 3. The same as in Fig. 2 but for sequence G. The gap is about 2.9.

rotations, and reflections. In practice, we put chain a over chain b by overlapping the two centers of mass, and then we find the optimal rotation of b which minimizes δ_{ab}. We pick the optimal angle from 1000 discrete values into which the $360°$ angle may be divided.

The distances between the native conformations and their lattice counterparts, as in Fig. 1, are found to be equal to 0.296 and ≈ 0 for sequences R' and G, respectively.

IV. THE FOLDING TEMPERATURE

We now proceed to the equilibrium characterization of the sequences introduced in this paper. The parameter that plays a primary role in determining the folding characteristics is the folding temperature T_f. In the case of lattice models, T_f may be defined [1] as the temperature at which the probability of occupying the native state becomes equal to $\frac{1}{2}$. For the off-lattice model, however, the native state has, strictly speaking, zero measure and one should deal with the probability of occupying the native valley.

One may define the folding temperature T_f through the following procedure. Suppose that δ is the distance to the native state and $\mathcal{P}(\delta)$ is the probability for a conformation to be in this distance away from the native state. Thus the probability to find the system in the immediate vicinity of the native state is given by
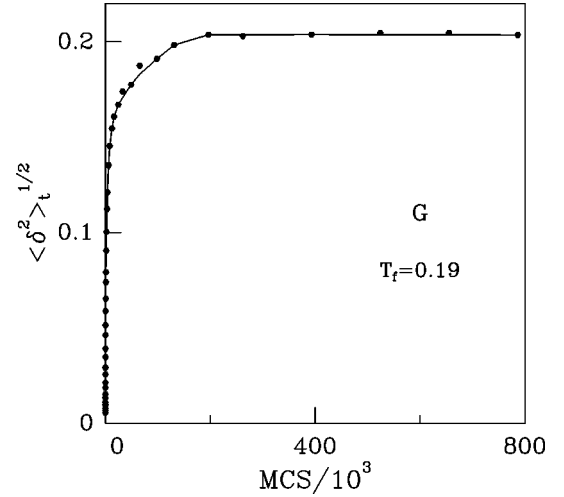

FIG. 4. The dependence of the distance to the native state on time measured in Monte Carlo steps (MCS) for sequence G at $T_f = 0.19$. In order to compute $\langle \delta^2 \rangle_t$ we update the monomer positions randomly within circles of radius of 0.01 (the choice of 0.02 yields similar results). The results are averaged over 400 Monte Carlo trajectories. The basin corresponding to the saturation value is $\delta_c \approx 0.2$.

$$\mathcal{P}_0 = \int_0^{\delta_c} \mathcal{P}(\delta) d\delta, \qquad (3)$$

where δ_c is a cutoff distance. The folding temperature T_f is then defined as the temperature at which $\mathcal{P}_0 = 1/2$.

The size of the native basin δ_c was estimated by the shape distortion approach [10]. In this approach one starts from the native state and performs random displacements of invidual beads in the chain, through a Monte Carlo routine. The distance to the native state is calculated by using Eq. (2) and the results are averaged over many Monte Carlo trajectories. Below some critical temperature which may be interpreted as a folding temperature T_f, the distance to the native state gets saturated at sufficiently long time scales. The saturation value of this distance at the critical temperature can serve as the size of the native basin δ_c. Figure 4 shows the dependence of $\langle \delta^2 \rangle_t^{1/2}$ on the Monte Carlo steps for sequence G and $T_f = 0.19$ above which no saturation regime is seen. It is found that $\delta_c \approx 0.2$ and $\delta_c \approx 0.09$ for G and R', respectively [10]. The corresponding values of the folding temperatures obtained in this way are equal to $T_f \approx 0.19$ and $T_f \approx 0.09$ for G and R' [10].

TABLE II. The x-y coordinates of the native conformation for R' and $k = 25$.

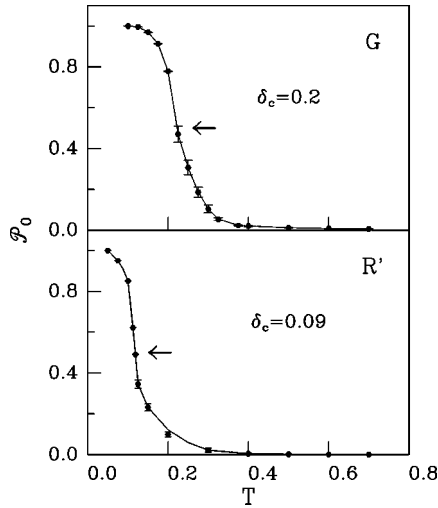| Monomer	x_N	y_N	Monomer	x_N	y_N
1	0	0	9	−0.037572	2.225111
2	1.005509	0.573347	10	−1.073118	2.718455
3	2.091357	0.126744	11	−0.992672	1.573907
4	2.009697	1.279681	12	−0.956188	0.436143
5	1.946778	2.424499	13	−0.972265	−0.723659
6	0.845883	2.806329	14	0.116648	−1.062284
7	0.997773	1.647486	15	1.123476	−0.470181
8	−0.035144	1.079819	16	2.154153	−0.984088

FIG. 5. The dependence of $\mathcal{P}_0$ on T for sequences G and R'. The arrow indicates the position of the folding temperature which is equal to $T_f = 0.23$ for sequence G and $T_f \approx 0.12$ for R'. The results are averaged over 10 – 40 Monte Carlo trajectories.

Figure 5 shows the probability of being in the native valley versus temperature for G and R'. $\mathcal{P}_0$ has been determined by estimating the probability of staying in the native valley by starting in the native state and monitoring the system for 5×10^6 Monte Carlo steps (the distance updating is of order 0.01). Doubling the length of the run does not affect the results in any visible manner. For sequences G and R', we have found $T_f \approx 0.23$ and $T_f \approx 0.12$. These values of T_f are close to those found by the shape distortion method [10].

Figure 6 shows the distribution $\mathcal{P}(\delta)$ at different temperatures for sequence R'. As the temperature increases the maximum becomes wider and moves toward larger distances to the native state δ.

We now turn to the temperature dependence of the folding time t_{fold}. At high temperatures, reaching the native state takes a long time due to entropic effects. At low temperatures, on the other hand, glassy phenomena may set in and make the folding process extremely slow. Thus t_{fold} plotted against T typically shows a minimum at a certain temperature $T_{\min}$. The idea of the existence of the glassy phase in proteins has originated in the Bryngelson and Wolynes stud-
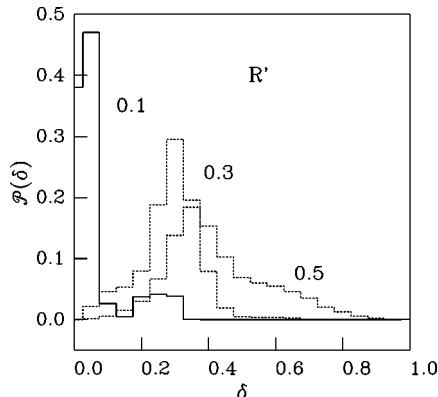

FIG. 6. The probability distribution function $\mathcal{P}(\delta)$ of finding a state which is distance δ away from the native state for sequence R'. The values of T are indicated next to the curves.
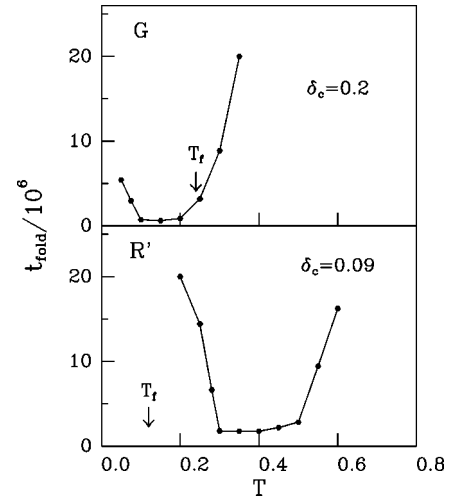

FIG. 7. The temperature dependence of folding time for sequences G and R'. The results are obtained using 20–50 starting configurations.

ies of the random energy model [7,12] and was subsequently tested in numerical simulations of lattice models by Socci and Onuchic [13]. Notice that T_f is a characteristic temperature that relates to equilibrium whereas $T_{\min}$ is a characteristic temperature that relates to dynamical properties. Even though at $T = T_{\min}$ folding is the fastest this temperature also marks the onset of the glassy effects because they become stronger and stronger on departing from $T_{\min}$ towards lower temperatures. Thus if T_f is significantly less than $T_{\min}$ a sequence is a bad folder. If T_f is comparable to $T_{\min}$, or preferably larger than $T_{\min}$, then the sequence is a good folder [13]. It should be noted that as a characteristic temperature that relates to dynamics one often uses the glass transition temperature T_g [1,13]. T_g, however, depends on a cutoff value of the time used in calculations. Our preference here is to use $T_{\min}$, instead of T_g, not only because its definition is unique and independent of the value of the cutoff but also because the two quantities contain the same physics: good folders are sequences in which glassy effects are not important around the folding temperature.

In order to obtain t_{fold} we start from random configurations and evolve them through the Monte Carlo (MC) process until the native state is reached. t_{fold} is defined as the median number of MC steps after which the system reaches the basin of the native state for the first time. The cutoff value of MC steps is taken to be 20×10^6. The number of starting configurations varies from 20 to 40 depending on T. Our results are presented in Fig. 7 for sequences R' and G. We have $T_{\min} = 0.4 \pm 0.1$ and $T_{\min} = 0.15 \pm 0.05$ for R' and G, respectively. Since $T_f = 0.12$ and $T_f = 0.23$ for these sequences one can see that R' is a bad folder whereas G is a good one. We emphasize that on-lattice Go and rank-ordering schemes lead to comparable ratios of $T_f / T_{\min}$.

V. THERMODYNAMIC PROPERTIES

We now proceed to a further characterization of the sequences by considering the thermodynamic quantities such as the specific heat C and the structural susceptibility χ [14]. It has been suggested [14–16] that a small temperature dif-

ference between the maximum in C and the maximum in χ when plotted against T indicates good folding properties. Thus studies of these two quantities may serve as a substitute for the information about the kinetics of folding.

In the case of an off-lattice model the departures of the sequence geometry from its native conformation may be described through the structural overlap function [3] as

$$\chi_s = 1 - \frac{1}{N^2 - 3N + 1} \sum_{i \neq j, j \pm 1} \Theta(\delta_c - |r_{ij} - r_{ij}^N|), \quad (4)$$

where r_{ij} is the distance between the beads i and j for a given conformation, r_{ij}^N is the corresponding distance in the native conformation, and $\Theta(x)$ is the Heaviside function. Here δ_c denotes the size of basin as defined in the preceding section. If $|r_{ij} - r_{ij}^N| \leq \delta_c$ the beads i and j are assumed to form a contact within the native valley. The susceptibility-like parameter χ is defined as the thermal fluctuation of χ_s:

$$\chi(T) = \langle \chi_s^2(T) \rangle - \langle \chi_s(T) \rangle^2, \quad (5)$$

where the angular brackets indicate a thermodynamic average.

The specific heat is defined in the usual way, i.e., by the energy fluctuations:

$$C = \frac{\langle E^2 \rangle - \langle E \rangle^2}{T^2}. \quad (6)$$

A peak in C may be interpreted as corresponding to the onset of slow kinetics and the peak in χ as corresponding to the folding temperature.

We calculate the thermodynamic quantities using the native state as the starting configuration. The results are averaged over many MC trajectories. The equilibration is checked by monitoring the stability of the data against at least three-times longer runs. We have used 5×10^6 MC steps as in the studies of $\mathcal{P}_0$ (the first 2.5×10^6 steps are not taken into account when averaging). This number of MCS appears to be enough to equilibrate the system in the temperature interval we use. We have also found that it is harder to reach the equilibrium using SAW's as starting configurations.

The temperature dependence of χ and C for sequences G and R' is shown in Fig. 8. The peak in χ almost coincides with T_f. For G maxima of χ and C are located at the same position, suggesting that sequence G is a good folder [14]. This result agrees with that obtained in the preceding section by studying kinetics of folding.

For sequence R', the peak in C is found to be broad and it is located at a substantially higher T than the peak in χ
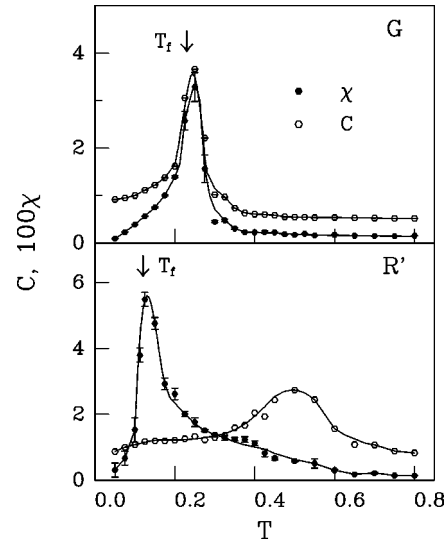

FIG. 8. The temperature dependence of χ and C for G and R'. The starting configurations are the native state. The results are averaged over ten MC trajectories. The arrows indicate the position of T_f found from the condition that the probability of staying in the native state is equal to 1/2.

confirming bad folding properties of R'. So studies of thermodynamics properties and folding kinetics lead to the same conclusion about folding in our model.

VI. CONCLUSIONS

In this work we have constructed and studied two off-lattice Lennard-Jones models in two dimensions. The long-range nature of interactions between amino acids in off-lattice models may lead to too much frustration, as in the case of sequence R', and thus to bad folding properties. Restricting the number of interactions only to the native contacts, as in sequence G, may reduce the frustration and bring about good folding properties. We have demonstrated this by studying both thermodynamic and dynamical properties of the two sequences. The values of the folding temperature found from the probability of getting out of the native state roughly agree with those obtained by the shape distortion approach. We have studied the kinetic of folding with a simplified MC dynamics whereas previous works have been focused on equilibrium aspects. Similar to the on-lattice models, both approaches give the same information about folding properties.

[1] K. A. Dill *et al.* Protein Sci. **4**, 561 (1995).

[2] A. Irback, C. Peterson, and F. Pottast, Proc. Natl. Acad. Sci. USA **93**, 9533 (1996); Phys. Rev. E **55**, 860 (1997); A. Irback, C. Peterson, F. Pottast, and O. Sommelius, J. Chem. Phys. **107**, 273 (1997).

[3] D. K. Klimov and D. Thirumalai, Phys. Rev. Lett. **79**, 317 (1997); T. Veitshans, D. K. Klimov, and D. Thirumalai, Folding Des. **2**, 1 (1997); e-print cond-mat/9611065.

[4] N. Go and H. Abe, Biopolymers **20**, 1013 (1981).

[5] I. Shrivastava, S. Vishveshwara, M. Cieplak, A. Maritan, and

J. R. Banavar, Proc. Natl. Acad. Sci. USA **92**, 9206 (1995); M. Cieplak, S. Vishveshwara, and J. R. Banavar, Phys. Rev. Lett. **77**, 3681 (1996).

[6] M. Cieplak and J. R. Banavar, Folding Des. **2**, 235 (1997).

[7] J. D. Bryngelson and P. G. Wolynes, Proc. Natl. Acad. Sci. USA **84**, 7524 (1987).

[8] G. Iori, E. Marinari, and G. Parisi, J. Phys. A **24**, 5349 (1991).

[9] M. V. Struglia, J. Phys. A **28**, 1469 (1995).

[10] M. S. Li and M. Cieplak (unpublished).

[11] M. Fukugita, D. Lancaster, and M. G. Mitchard, J. Phys. A **25**, L121 (1992).

[12] J. D. Bryngelson and P. G. Wolynes, J. Chem. Phys. **93**, 6902 (1989).

[13] N. D. Socci and J. N. Onuchic, J. Chem. Phys. **101**, 1519 (1994).

[14] C. J. Camacho and D. Thirumalai, Proc. Natl. Acad. Sci. USA **90**, 6369 (1993).

[15] C. J. Camacho and D. Thirumalai, Europhys. Lett. **35**, 627 (1996); D. K. Klimov and D. Thirumalai, Phys. Rev. Lett. **76**, 4070 (1996).

[16] D. Thirumalai, J. Phys. I **5**, 1457 (1995).