ELSEVIER

Biophysical Chemistry 106 (2003) 81–89

Biophysical Chemistry

www.elsevier.com/locate/bpc

3D-Lattice Monte Carlo simulations of model proteins. Size effects on folding thermodynamics and kinetics

K. Leonhard[a], J.M. Prausnitz[a,b], C.J. Radke[a,c,*]

[a]*Department of Chemical Engineering, University of California, Berkeley, CA 94720-1462, USA*
[b]*Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[c]*Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

Received 13 May 2003; received in revised form 20 June 2003; accepted 20 June 2003

Abstract

Recently, we devised an energy scale to vary systematically amino-acid residue–solvent interactions for Monte Carlo simulations of lattice-model proteins in water. For 27-mer proteins, the folding behavior varies appreciably with the choice of interaction parameters. We now perform similar simulations with 64-mers to study the size dependence of the optimal energy parameter set for representing realistic behavior typical of many real proteins (i.e. fast folding and high cooperativity for single chains). We find that 64-mers are considerably more stable and more cooperative compared to 27-mers. The optimal interfacial-interaction-energy parameter set, however, is relatively size independent.

Keywords: Lattice simulation; Interaction energies; Protein folding; Solvation; Size effects

1. Introduction

Monte Carlo (MC) simulations of folding–unfolding events of lattice proteins have been widely pursued in the literature. Several studies have been performed with different-size chains. Li et al. [1] enumerated all structures of compact 27-mers for all sequences of a 2-letter hydrophobic-polar (HP) model to find out how many sequences have unique ground states and which structures have a high designability, i.e. those structures with a unique ground state of many sequences. Du et al. [2] used a 48-bead homopolymer and a 48-bead Go model [3] to investigate effects of interaction potential specificity. Shakhnovich et al. [4–6] studied 27-mers with a 2-letter model as well as 36-mers (with emphasis on folding entropy and energy) and 80-mers (for correlation of stability and kinetic accessibility) with the 20-letter Miyazawa–Jernigan (MJ) model. Tiana et al. [7] studied the effect of mutations on the folding of a 36-bead model protein interaction via MJ energies. Pande and Rokhsar [8] employed a 48-bead lattice chain and Go interactions to investigate the existence of a folding pathway ensemble in contrast to a classical single pathway. Mirny et al. [9] studied sequence effects on folding speed and

*Corresponding author. Tel.: +1-510-642-5204; fax: +1-510-642-4778.
E-mail address: radke@cchem.berkeley.edu (C.J. Radke).

0301-4622/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0301-4622(03)00185-6

stability of a 48-mer using MJ interaction energies. Crippen and Chhajer [10] investigated the effect of the interaction parameters in an HP model on the folding of chains in 2D, allowing for non-compact native states. They also considered very short chains, eight to ten segments long, in 3D.

One study has been performed with a 125-bead lattice protein by Dinner et al. [11]. The interactions between amino acids were drawn from two Gaussian distributions, one with an attractive mean for native contacts and one with a neutral mean for non-native contacts. Hence, the lattice protein of Dinner et al. is described by a 125-letter energy scale, in contrast to the 20 amino acids that are used for building proteins in nature. Unlike smaller chains, strict thermodynamic and kinetic requirements have to be fulfilled to make a 125-mer behave protein-like. Namely, the protein core has to be relatively stable and the periphery has to be less stable. Only then, when the core forms first and cooperative contacts [11] are formed will the 125-mer fold to a native form (i.e. the sequence is well designed).

Recently, we generated amino acid–solvent interactions that are based on the MJ interaction energy scale for amino acid interresidue contacts [12], but that allow us, by changing one parameter, $\bar{\omega}$, to alter the trend of lattice proteins to prefer compact structures [13]. In addition, with a second parameter, C_s, our interaction energy-scale allows us to adjust the difference between two interaction energies, viz., the interaction energy between a hydrophobic amino acid and the solvent and the interaction energy between a hydrophilic amino acid and the solvent. In other words, the amino acid–solvent interactions can be altered independently from the amino acid–amino acid interactions, which we take from Miyazawa and Jernigan [12]. Consequently, we can change the relative strength of general effects like the hydrophobicity of the amino acids and the contrast of their hydrophobicity compared to the specific amino acid–amino acid interactions. Using C_s and $\bar{\omega}$, we can design solvents of different polarity, study protein folding and aggregation in these solvents, and then, upon comparing with experimentally observed behavior for aqueous protein solutions, decide which of them is most conducive to realistic protein behavior [13].

27-mers are widely used in folding studies because they are known to reproduce generic effects of folding, such as the folding funnel and the hydrophobic collapse [14]. Our previous study [13] has shown that the folding time and the cooperativity for a 27-mer depend strongly on the solvent interactions and, therefore, on C_s and $\bar{\omega}$. However, the energy scale that is optimal for 27-mer behavior is different from the original MJ interaction energies, which are based on proteins approximately 100–400 amino acids long. Therefore, we want to investigate in this work whether this difference is due to the small size and the large surface volume ratio of the 27-mer compared to proteins used to estimate the original MJ potential or whether the optimum in energy parameter space is universal for lattice proteins. Therefore, we now investigate for a variety of energy parameters, the effect of chainlength on the behavior of a single chain.

2. Interaction energy

Previously, we proposed the following expression for the amino acid–solvent interaction energy, e_{i0} [13]:

$$e_{i0} = \frac{1}{2}(1 - C_s)e_{ii} + \bar{\omega} + \frac{C_s}{2n}\sum_{i=1}^{n} e_{ii} \qquad (1)$$

where the amino acid–amino acid self interaction energy, e_{ii}, is taken from Table V of Miyazawa and Jernigan [12], as are all other non-self interaction energies, e_{ij}. C_s in Eq. (1) is a contrast parameter. The closer is C_s to 0, the solvent more equally solvates both hydrophobic and hydrophilic amino acids. The more positive is C_s, the more favorable are the contacts between the solvent and hydrophilic amino acids. Parameter $\bar{\omega}$ describes the average interaction of an amino acid with the solvent. When $\bar{\omega}$ is negative, amino acids, on average, have attractive interactions with the solvent. When $\bar{\omega}$ is positive, they repel the solvent and hence effectively attract each other. Because it is more convenient to think in terms of amino acid interactions than in terms of amino acid–

solvent interactions, we call $\bar{\omega}$ the 'effective amino acid–amino acid mean attraction' or briefly 'mean attraction'. Although based on exchange energies, a positive value of $\bar{\omega}$ in this context means an effective attraction between residues.

3. Simulation technique

Details of our simulations are available elsewhere [13]. Here, we give only a brief overview of the simulation techniques. All simulations are performed on a 3D-cubic lattice with periodic boundary conditions and a size of $20\times20\times20$ sites. Proteins are modeled by self-avoiding chains of beads, each bead representing a Kuhn segment [15]. All sites that are not occupied by a chain group are occupied by an effective solvent molecule (i.e. an effective solvent molecule is a lattice site of the same size as a Kuhn segment and can represent a group of solvent molecules that corresponds to the size of the segment) [12].

The system energy is defined as the sum of the interaction energies over all pairs of adjacent sites on the lattice, but which are not neighbors on a chain. Accordingly, the system energy is described by the Ising-like Hamiltonian:

$$H=\sum_{i=1}^{L}\sum_{j=1}^{L}\sum_{k=1}^{L} e_{(i,j,k),(i+1,j,k)}+e_{(i,j,k),(i,j+1,k)} + e_{(i,j,k),(i,j,k+1)} \quad (2)$$

where L is the size of the lattice in lattice units, and $e_{(i,j,k),(l,m,n)}$ is the interaction energy between an amino acid or solvent bead on site (i, j, k) and one on site (l, m, n). Only nearest-neighbor interactions are non-zero. During the simulation, end-flip, kink-bend and crankshaft MC moves are attempted. If we try to move a chain bead to a site occupied by another chain bead, the move is immediately rejected because the chains are self-avoiding. Otherwise, an exchange of the protein bead with a solvent bead on this site is attempted. Then, the move is accepted or rejected depending on the energy change ΔE associated with the move in accordance with the Metropolis criterion:

$$\Delta E\leqslant 0\text{: always accept,} \quad (3)$$

Table 1
One-letter code for the sequences of amino acids used in the simulations

C_s	Sequence
0.1	KALAAGGVTTAVSDLCSSVVPDLIRITDMGGW KGAEAIKYLVHELSKSTGNDKQRFPNNYFQPE
0.2	KEKSTAGRVASGVLDSVACGVLGDIDTLQGSP IAKLKTFYGNKFNDVEASQAHMIRWPNYTLPE
0.4	KDKDTASQVAKGVLPNVATGVLGEIDALRSSP IARLKGFYGSKCTEFDASGVHMIQYTNPNLWE

and

$$\Delta E>0\text{: accept with probability } \exp\left(\frac{-\Delta E}{k_{\mathrm{B}}T}\right) \quad (4)$$

where k_{B}, the Boltzmann constant, has the value of one MJ-energy unit per one associated temperature unit. Because MJ energies are based on contact frequencies, they are only relative energies. To our knowledge, no conversion to conventional energy units is available. Therefore, for our purpose here, temperature can be expressed in arbitrary units.

The types of the 64 amino acid beads for a chain are chosen to reflect a composition corresponding to a typical real protein. Then the sequence is obtained by simulated annealing with the result shown in Table 1. The concentration of the chains in our simulation box is approximately 0.8 vol% for a single 64-mer.

In a MC simulation, free energy differences between states can be observed by counting the number of occurrences of the states. We use the number of native contacts, N_{nat}, called order parameter, to create a histogram for the chain in a simulation and we count how often the chain is in each state corresponding to the order parameter. After the simulation, the relative free energy $F(\mathrm{A})$ of state A can be calculated from the number of times the chain was found in the state $Z(\mathrm{A})$.

$$F(\mathrm{A})=-k_{\mathrm{B}}T\ln Z(\mathrm{A}) \quad (5)$$

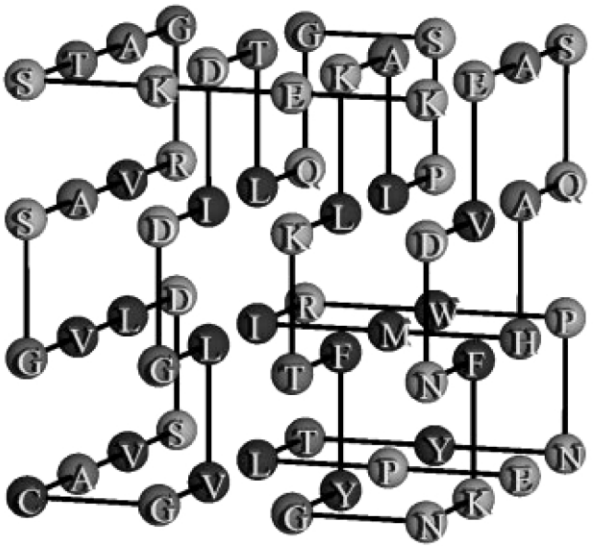

Fig. 1. Schematic of the native state of the fastest folding 64-mer ($C_s = 0.2$). Light gray is for hydrophilic amino acid beads, medium gray for neutral and dark gray for hydrophobic beads. Each bead is labeled with the one-letter abbreviation for the amino acid it represents (see Figure 2 of [13] for the corresponding 27-mer).

4. Results

Enlarging the protein size from the small 27-mer to a larger 64-mer leads mainly to quantitative differences. Fig. 1 shows the native fastest-folding conformation that we use for our 64-mer simulations while Table 2 summarizes results for the midpoint temperature, T_m, folding time, t_f, at T_m, total number of contacts, N_{tot}, at T_m and cooperativity, $\Delta T/T_m$, for the 64-mer single chain. Comparative results for a 27-mer are reproduced in Table 3 [13]. First, the folding and refolding speed is significantly slower for the longer chain simply because of the larger conformational space. The mean folding time t_F is the average number of MC time steps for one folding/unfolding cycle, where the chain is considered folded when it has the maximal number of native contacts (81) and unfolded when it has 20 or fewer native contacts. Overall stability and cooperativity are considerably larger for the long chain. Crippen and Chhajer [10] wrote a short discussion of different definitions of cooperativity in one of their papers. We use a slightly different definition for the cooperativity, $\Delta T/T_m$, where the cooperativity is calculated from the temperatures where on average 0.9, 0.5 and 0.1 times the maximum number of native contacts can be observed:

$$\frac{\Delta T}{T_m} = \frac{T(N_{nat} = 0.1N_{nat,max}) - T(N_{nat} = 0.9N_{nat,max})}{T(N_{nat} = 0.5N_{nat,max})}. \quad (6)$$

Table 2
64-mer midpoint temperature T_m, mean folding time t_F at T_m, number of total contacts N_{tot} at T_m and cooperativity $\Delta T/T$ for the interaction energy parameter sets studied

C_s	$\bar{\omega}$	T_m	t_F per 10^8 steps	N_{tot}	$\Delta T/T_m$
0.10	0.06	0.29	50	48.9	1.29[a]
0.10	0.16	0.40	1.6	58.8	1.02
0.10	0.26	0.48	3.1	63.7	1.09
0.20	0.06	0.31	3.0	50.0	[b]
0.20	0.16	0.43	0.5	55.8	0.73
0.20	0.26	0.53	1.5	62.9	0.75
0.40	0.22	0.45	10	59.9	[b]
0.40	0.32	0.56	3.7	62.0	1.32
0.40	0.42	0.60	33	66.9	1.38

[a] The maximum compact state is not the most stable state, but there is a temperature where on average the number of native contacts is 90% of the maximum of native contacts.

[b] The average number of native contacts is less than 90% of the maximum number for all temperatures.

Table 3
27-mer midpoint temperature T_m, mean folding time t_F at T_m, number of total contacts N_{tot} at T_m and cooperativity $\Delta T/T$ for the interaction energy parameter sets studied

C_s	$\bar{\omega}$	T_m	t_F per 10^6 steps	N_{tot}	$\Delta T/T_m$
0.10	0.06	0.27	0.9	17.7	1.76
0.10	0.16	0.37	1.2	19.5	1.39
0.10	0.26	0.46	1.9	21.1	1.45
0.20	0.06	0.30	0.5	17.8	1.26[a]
0.20	0.16	0.41	0.8	18.9	1.05
0.20	0.26	0.50	1.4	20.8	1.07
0.40	0.22	0.43	3.5	19.7	1.91
0.40	0.32	0.53	3.1	20.7	1.56
0.40	0.42	0.61	4.1	21.6	1.59

[a] The maximum compact state is not the most stable state, but there is a temperature where on average the number of native contacts is 90% of the maximum.

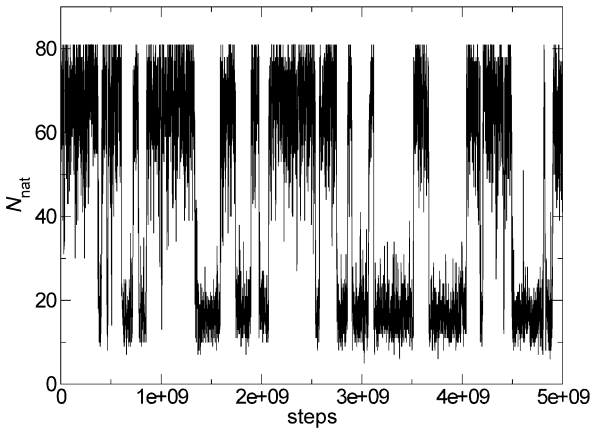

Fig. 2. Folding and unfolding kinetics of the 64-mer as measured by the number of native contacts at T_m. $C_s = 0.2$ and $\bar{\omega} = 0.26$.

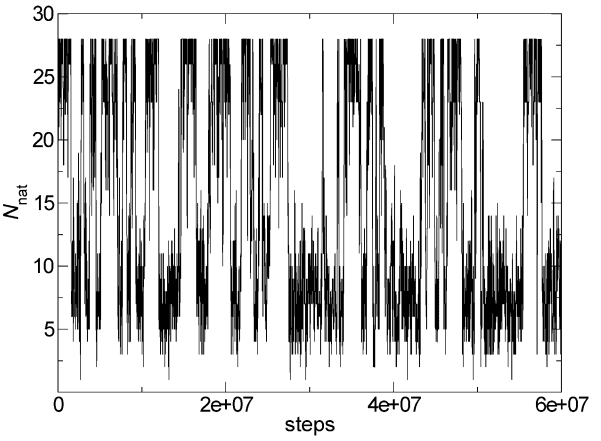

Fig. 3. Folding and unfolding kinetics of the 27-mer as measured by the number of native contacts at T_m. $C_s = 0.2$ and $\bar{\omega} = 0.26$.

A low value of $\Delta T/T_m$ corresponds to a high cooperativity and vice versa. The number of native contacts, N_{nat}, is the number of those contacts that are also present in the native state. The increased stability is indicated by the midpoint temperature T_m, which is defined as the temperature where, on average, half of the native contacts are present during a simulation. For a 64-mer with a $4\times4\times4$ cube as native state having 81 native contacts, this means that T_m is the temperature where $N_{nat} = 40.5$, on average. For the 27-mer, T_m is the temperature where $N_{nat} = 14$, on average. For each interaction energy set, for the 64-mer, T_m is 5–10% larger than that for the 27-mer. This difference is attributed to the larger hydrophobic core and to the larger number of specific contacts of the larger molecule. The 27-mer exhibits only one group with no water contacts, but the 64-mer displays eight, although the ratio of the total number of sites in the two *n*-mers is only approximately 2.4.

In Fig. 2, the number of native contacts, N_{nat}, is plotted vs. the number of MC steps for a relatively strong mean attraction, $\bar{\omega} = 0.26$. This figure shows that the compactness of the folded and the unfolded states is much more separated (unfolded state: $\sim$5–25 native contacts and folded state: $\sim$50–82 native contacts) for the longer 64-mer protein chain than for the 27-mer shown in Fig. 3. In addition, for the 64-mer, the plot of the N_{nat} vs. T is much steeper near T_m (data not shown). Compared to the 27-mer model protein, these observations reflect a stronger cooperativity.

For a weak mean attraction, $\bar{\omega} = 0.06$, in Fig. 4, only the core of the 64-mer is folded. A substantial portion of the corona part of the coil remains unfolded. As a consequence, the completely folded state is visited only rarely. Therefore, when the mean attraction is weak, the refolding rate is reduced. Because of this effect for a weak mean

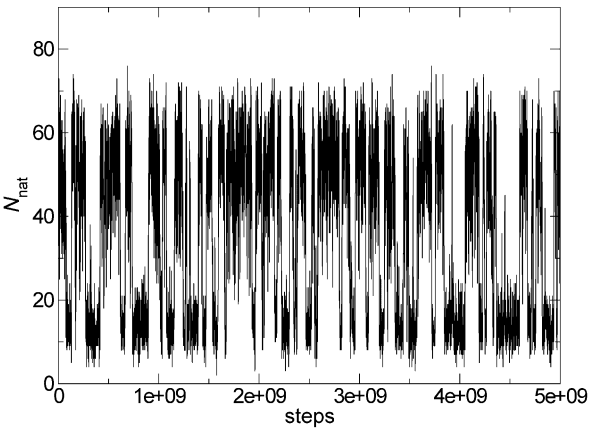

Fig. 4. Folding and unfolding kinetics of the 64-mer as measured by the number of native contacts at $1.03T_m$. $C_s = 0.2$ and $\bar{\omega} = 0.06$.

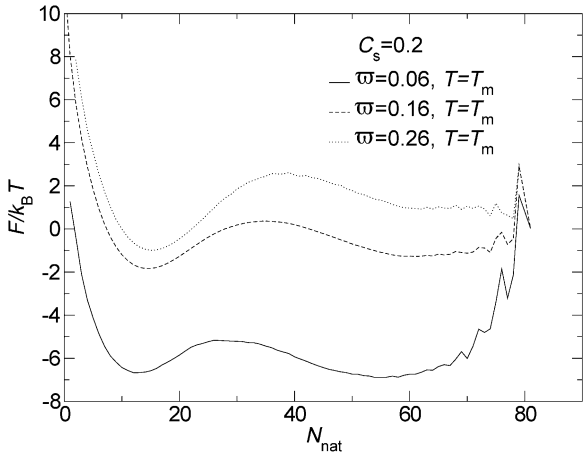

Fig. 5. Reduced-free-energy profile of the folding–unfolding equilibrium of the 64-mer. $C_s = 0.2$.

attraction, fastest refolding occurs at slightly higher values of $\bar{\omega}$ than it does for the 27-mer (for $C_s < 0.4$). The optimal value of C_s, however, remains at 0.2, as for the short protein [13]. For the 64-mer, fastest folding and highest cooperativity occur at the same mean attraction. With the 27-mer, however, fastest folding occurs at a smaller $\bar{\omega}$ than the highest cooperativity. Because the number of chain segments and the surface–volume ratio of the 64-mer are closer to those of real proteins than the values for the 27-mer, the difference between the two optima for the small protein is attributed to the small size of the 27-mer. For a weak mean attraction, fastest refolding for 64-mers takes place at a temperature somewhat below T_m. For stronger mean attractions, the chain folds fastest at T_m, as observed for the 27-mer [13].

Table 2 for the 64-mer also shows the number of total contacts N_{tot} at T_m, which is the number of contacts between the beads of the chain and all other amino acid beads, omitting contacts between the two neighbors that are consecutive on the chain (one neighbor for a terminal bead). A contact between two beads exists when the beads are nearest neighbors on the lattice. N_{tot} increases for every C_s when $\bar{\omega}$ increases, although the number of native contacts remains 40.5, by definition. This indicates that the coil is more compact and that more non-native contacts are formed for a high mean attraction, or in other words, in this case the interactions are less specific, in agreement with our results for the 27-mers.

Fig. 5 shows the free-energy landscape for the folding/unfolding process of the 64-mer. This figure can be compared with the results for the 27-mer in Fig. 6 [13]. The profiles are plotted vs. the number of native contacts N_{nat}. The three lines in these figures represent different mean attractions, all at the respective midpoint temperatures. When comparing the free-energy profile for the 64-mer to that for the 27-mer, several features merit attention that are different for the two chains. First, we examine the peak at 79 native contacts in the 64-mer (a state with one native contact less than the maximum number of native contacts, which is 81 for the 64-mer, does not exist for a self-avoiding chain on a cubic lattice). From the folding kinetics, we find that usually the last step to form 81 native contacts is one of three possible crankshaft moves at one of the edges of the cube (either amino acids VGSA, FNDV or PSGQ in Fig. 1 are involved). Such a move creates four native contacts in one step, indicating that the system can 'jump' from 77 to 81 contacts. The states with 78 or 79 native contacts, which have a very high free energy, provide only alternative paths to the native state. Therefore, the peak at

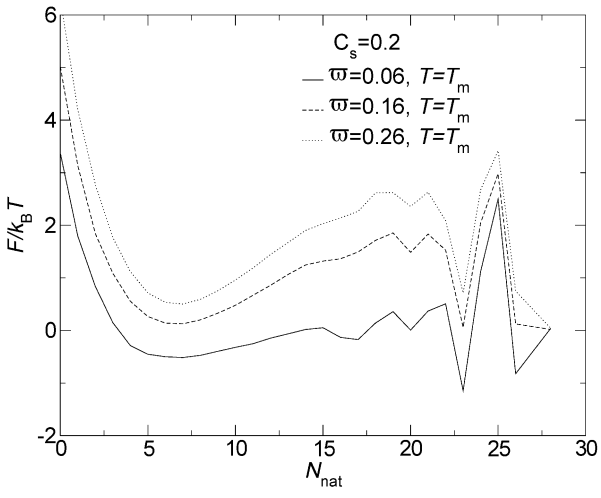

Fig. 6. Reduced-free-energy profile of the folding–unfolding equilibrium of the 27-mer. $C_s = 0.2$.

$N_{nat}=79$ in Fig. 5 is not a folding barrier; the 64-mer is a true 2-state folder. Conversely, the shorter 27-mer is a 3-state folder as can be seen by examining the free energies in Fig. 6 near $N_{nat}=23$ and 26 [13].

The actual folding barrier for the 64-mer is at approximately 20–40 native contacts, or 25–50% of the maximum number of native contacts in contrast to the 27-mer, where the barrier is at approximately the same absolute number of native contacts, 20, which is 70% of the total number. This difference in relative position of the folding barrier indicates that, for the 27-mer at midpoint temperature, the loss in entropy is stronger than the gain in stabilization energy for a large portion of the folding pathway until the folded state is almost reached. On the other hand, for the 64-mer and starting from the unfolded state, only relative few amino acids have to be put into their places against a large entropy loss, and then the rest fall into place almost automatically, until the free-energy minimum of the folded state is reached. This observation agrees nicely with the higher cooperativity that we find for the 64-mer and also agrees with the results of Dinner et al. [11] for a 125-mer.

For a weak mean attraction, $\bar{\omega}=0.06$, a configuration of the 64-mer with 70 of the native contacts of the maximal compact state is the lowest energy state. However, the lowest free-energy state at T_m is at approximately 55 native contacts due to entropy effects. The lowest energy state of this protein is indicated in Fig. 7. In this conformation, four pairs of beads have performed crankshaft moves to rotate out and to allow water access a larger number of hydrophilic beads relative to that in the maximal compact state. The amino acids involved in this change from the maximal compact state are QGSP, KNGY, VGCA and VGSA. They can be located in that order when looking clockwise around the original cube, starting in the upper left corner. Because this lowest energy state is only partially folded, the free-energy minimum that corresponds to the folded state is rather broad at approximately 50–70 native contacts while the fully-folded state has a substantially higher free energy. Therefore, for this parameter set, our 64-mer chain has much flexibility in its free-energy basin corresponding to the native state at the midpoint temperature, and even at a much lower temperature. For $\bar{\omega}=0.16$, the $4\times4\times4$ cube is the lowest energy state, but it is only marginally stable compared to states with 77 native contacts. Since the states with 77 native contacts are entropically favored, they have a lower free energy than the fully-folded state at T_m and at somewhat lower temperatures. Therefore, it seems most reasonable to consider an ensemble of states as native, as suggested by Crippen and Chhajer [10], not a single state for these conditions. In our work, reported values for the folding time and the cooperativity are always based on the fully-folded state. In a follow-on paper, we investigate the effect of different choices of the definition of the native state on the interpretation of 64-mers aggregation.

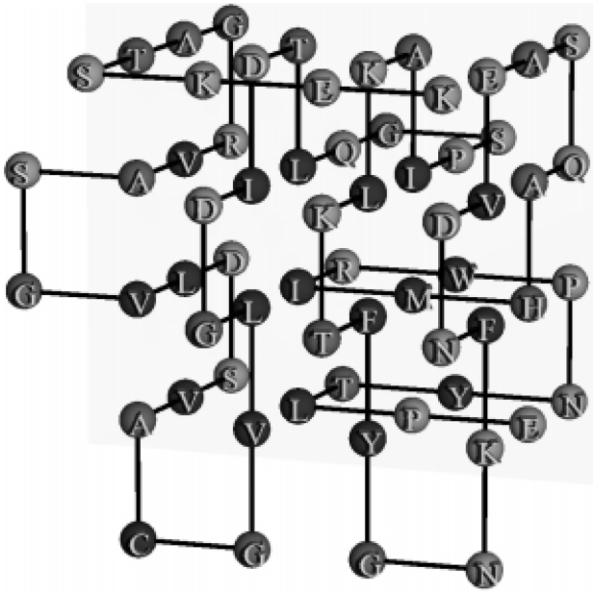

Fig. 7. Schematic of the native state of a non-compact 64-mer ($C_s=0.2$ and $\bar{\omega}=0.06$). Light gray is for hydrophilic amino acid beads, medium gray for neutral and dark gray for hydrophobic beads. Each bead is labeled with the one-letter abbreviation for the represented amino acid.

5. Discussion

In our study for 27-mers, we divided the energy parameter space of C_s and $\bar{\omega}$ into four regions [13]

(I–IV) of decreasing mean attraction. In Regions I and II, we find compact native states and in Regions III and IV there are non-compact structures. Regions II and III feature the most protein-like behavior of our lattice chains and they exist only when the contrast parameter C_s is between approximately 0.05 and 0.35. In Region I, the mean attraction is too strong for fast folding and high cooperativity, and in Region IV it is too weak to cause a folding–unfolding transition. The same four regions with approximately the same boundaries apply here for 64-mers. The energy-parameter set found to be optimal for cooperativity for the 27-mer ($\bar{\omega}=0.16$ and $C_s=0.2$) is also optimal for high cooperativity for the 64-mer. The optimal parameter set is located in Region II, close to the border to Region III. For the 27-mer, there is a lower mean attraction than for highest cooperativity that features fastest folding. However, for the 64-mer, the same parameter set is optimal for both. For the 27-mer, the folded state at fastest folding is scattered and barely separated from the unfolded state. We cannot speak of protein-like behavior in this case. Nevertheless, the fully-folded state is visited very frequently for such a 27-mer, because the free energy of the fully-folded state is not significantly higher than that for the semi-folded state. For the 64-mer, however, the folding speed starts to decrease shortly after the mean attraction is lowered sufficiently to make a slightly flexible native state.

6. Conclusions

It is known that natural proteins have a higher cooperativity than proteins in lattice-based models [16]. Therefore, the higher cooperativity and the higher stability of the 64-mer compared to the 27-mer make the 64-mer a more realistic protein model than the 27-mer. Higher cooperativity and stability are presumably caused by the larger hydrophobic core and the higher volume-to-surface ratio of the 64-mer.

We establish that the optimal parameters of our modified MJ interaction-energy scale do not change much, or not at all, with protein size, depending on the property considered. Therefore, the original MJ interaction energies seem to overestimate the amino acid hydrophobicity, as already noted by Miyazawa and Jernigan [12].

Acknowledgments

The authors thank Dusan Bratko for helpful discussions. For financial support the authors are grateful to DAAD (German Academic Exchange Service) for a fellowship to K.L. and to the Office for Basic Sciences of the US Dept. of Energy.

References

[1] H. Li, R. Helling, C. Tang, N. Wingreen, Emergence of preferred structures in a simple model of protein folding, Science 273 (1996) 666–669.
[2] R. Du, V.S. Pande, A.Y. Grosberg, T. Tanaka, E. Shakhnovich, On the role of conformational geometry in protein folding, J. Chem. Phys. 111 (1999) 10375–10380.
[3] N. Go, Theoretical studies of protein folding, Annu. Rev. Biophys. Bioeng. 12 (1983) 183.
[4] E.I. Shakhnovich, A.M. Gutin, Engineering of stable and fast-folding sequences of model proteins, Proc. Natl. Acad. Sci. 90 (1993) 7195–7199.
[5] E.I. Shakhnovich, Proteins with selected sequences fold into unique native conformation, Phys. Rev. Lett. 72 (1994) 3907–3910.
[6] V.I. Abkevich, A.M. Gutin, E.I. Shakhnovich, Free energy landscape for protein folding kinetics: intermediates, traps, and multiple pathways in theory and lattice model simulations, J. Chem. Phys. 101 (1994) 6052–6062.
[7] G. Tiana, R.A. Broglia, H.E. Roman, E. Vigezzi, E. Shakhnovich, Folding and misfolding of designed proteinlike chains with mutations, J. Chem. Phys. 108 (1998) 757–761.
[8] D.S. Pande, S. Vijay, Rokhsar, Folding pathway of a lattice model for proteins, Proc. Natl. Acad. Sci. 96 (1999) 1273–1278.
[9] L.A. Mirny, V.I. Abkevich, E.I. Shakhnovich, How evolution makes proteins fold quickly, Proc. Natl. Acad. Sci. 95 (1998) 4976–4981.
[10] G.M. Crippen, M. Chhajer, Lattice models of protein folding permitting disordered native states, J. Chem. Phys. 116 (2002) 2261–2268.
[11] A.R. Dinner, A. Šali, M. Karplus, The folding mechanism of larger model proteins: role of native structure, Proc. Natl. Acad. Sci. 93 (1996) 8356–8361.
[12] S. Miyazawa, R.L. Jernigan, Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation, Macromolecules 18 (1985) 534–552.
[13] K. Leonhard, J.M. Prausnitz, C.J. Radke, Solvent amino acid interaction energies in 3-D-lattice Monte Carlo

simulations of a model 27-mer protein. Folding thermodynamics and kinetics, Protein Sci., submitted for publication.
[14] M. Gruebele, The fast protein folding problem, Annu. Rev. Phys. Chem. 50 (1999) 485–516.
[15] R.H. Boyd, P.J. Phillips, The Science of Polymer Molecules, Cambridge University Press, 1993.
[16] H. Kaya, H.S. Chan, Polymer principles of protein calorimetry two-state cooperativity, Proteins: Struct. Func. Gen. 40 (2000) 637–661.