PROTEINS: Structure, Function, and Genetics 34:472–483 (1999)

The Folding Funnel Landscape for the Peptide Met-Enkephalin

Ulrich H.E. Hansmann,[1]* Yuko Okamoto,[2] and Jose N. Onuchic[3]
[1]*Department of Physics, Michigan Technological University, Houghton, Michigan*
[2]*Department of Theoretical Studies, Institute for Molecular Science, Okazaki, Japan*
[3]*Department of Physics, University of California at San Diego, La Jolla, California*

ABSTRACT **We study the free energy landscape of the small peptide Met-enkephalin. Our data were obtained from a generalized-ensemble Monte Carlo simulation taking the interactions among all atoms into account. We show that the free energy landscape resembles that of a funnel, indicating that this peptide is a good folder. Our work demonstrates that the energy landscape picture and folding concept, developed in the context of simplified protein models, can also be used to describe the folding in more realistic models. Proteins 1999;34:472–483.**

Key words: protein folding; energy landscape theory; generalized-ensembles; Monte Carlo simulations

INTRODUCTION

It is well known that a large class of proteins fold spontaneously into unique, globular shape.[1] However, the mechanism of protein folding has remained elusive. To describe many biochemical processes it may be sufficient to assert that folding occurs on a time scale no slower than protein biosynthesis, and that the information required to find the precise three-dimensional shape is contained in the one-dimensional sequence of the molecule. This simple description may not be sufficient if the prediction of protein structure from sequence and the design of truly novel protein-like molecules are to be achieved. In order to answer the practical questions of structure prediction and design, it seems that one must go a considerable distance beyond this phenomenology—a new viewpoint may be required. Such a new viewpoint is emerging from the analytical and numerical studies of minimal protein models by several groups (see for example refs. 2–7). Its framework is provided by energy landscape theory and the funnel concept, which assert that a full understanding of the folding process requires a global overview of the landscape. The folding landscape of a protein resembles a partially-rough funnel riddled with traps where the protein can transiently reside. There is no unique pathway but a multiplicity of convergent folding routes towards the native state.[5–7]

The importance of a funnel landscape can be seen by contrasting random heteropolymeric molecules and proteins. Both random heteropolymers and proteins have an underlying driving force to collapse, and for both molecules, the various competing interactions within the molecule and between the molecule and the surrounding solvent lead to a rugged energy landscape. However, unlike random heteropolymers, proteins adopt well-defined three-dimensional structures because there is a sufficient overall slope of the energy landscape so that the numerous valleys flow in a funnel toward the native structure.[2,6,8–10]

The essence of the funnel landscape idea is competition between the tendency towards the folded state and trapping due to ruggedness of the landscape. This competition is measured by the ratio between the folding temperature (T_f) and the glass temperature (T_g). Good folding protein sequences fold rapidly on minimally frustrated landscapes with large values of T_f/T_g.[5–7] Minimally frustrated sequences not only fold fast at relevant temperatures but are also robust folders, and therefore only weakly dependent on minor variations of the folding environment or to mutations. Energy landscape theory suggests a diversity of folding scenarios that have been explored by computer simulations of minimalist models (see for example references in[2,7,11–17]) and connections to studies of real proteins were proposed.[9,18]

A quantitative understanding of how the general parameters of the landscape relate to particular properties of a protein's sequence or final folded topology requires detailed molecular (and therefore less coarse-grained) calculations of folding free energy landscapes. Unfortunately, simulations of more realistic models of proteins where the interactions among all atoms in a protein are taken into account have been notoriously difficult (see for example ref. 19 for a review). These studies involve extremely intensive numerical calculations, and thus the number and size of systems that can be explored is limited. In addition, it is difficult for such simulations to provide a direct view of protein folding dynamics. Because of the rough energy landscape, simulations based on canonical Monte Carlo or molecular dynamics techniques will get trapped at low temperatures in one of the multitude of local minima separated by high energy barriers. Hence, there is always the danger that only small parts of

Grant sponsor: Japanese Ministry of Education, Science, Sports and Culture; Grant sponsor: National Science Foundation; Grant number: 96-03839; Grant sponsor: Research Excellence Fund, State of Michigan; Grant number: E 27448.
*Correspondence to: Ulrich H.E. Hansmann, Department of Physics, Michigan Technological University, Houghton, MI 49931-1295. E-mail: hansmann@mtu.edu
Received 26 June 1998; Accepted 29 October 1998

configuration space are sampled and physical quantities cannot be calculated accurately.

New numerical approaches have been developed to deal with these sampling difficulties. Generalized-ensemble techniques (for a recent review, see, for instance, ref. 20) like multicanonical algorithms[21] and simulated tempering[22] allow an efficient sampling of low-energy configurations, and calculation of accurate low-temperature thermodynamic quantities became feasible. The first application of one of these techniques to the protein-folding problem is given in ref. 23. By comparing the approach with recent experiments, it's usefulness was extensively tested and demonstrated in a study of helix-coil transitions of homo-oligomers of nonpolar amino acids[24,25] and, more recently, of the C-peptide of ribonuclease A.[26,27] A numerical comparison of three different generalized-ensemble algorithms can be found in ref. 28.

Even with these new and sophisticated techniques, there are still many concerns and difficulties in fully exploring the energy landscape of even medium-sized proteins, and for this reason we have restricted the analyses to only small peptides in this paper.

In general, such peptides are too small to fold and therefore not useful as a model for the folding process in larger proteins. However, there are a few small peptides known that exist at low temperatures in stable defined conformations. For one of these peptides the characteristic temperatures of folding were recently determined numerically in ref. 29. These temperatures were shifted to lower values than expected for larger proteins, but otherwise the results of that study indirectly support the energy landscape picture and funnel concept. Here, we re-examine our data to directly investigate the free energy landscape of the peptide. In that respect, our work is similar to the earlier work in ref. 30 but uses this new sampling technique.

Our system of choice is the linear peptide Met-enkephalin which has the amino-acid sequence Tyr-Gly-Gly-Phe-Met. The experimental studies of this penta-peptide[31–33] were mainly motivated by its biological significance as a neurotransmitter. Compact and defined structures were observed for membrane bound molecules,[32] in crystals[31] or organic solvents. However, studies of this peptide in aqueous solution indicate that it exists at room temperature as an ensemble of extended coil structures with low frequency of folded structures.[33]

Since the energy landscape is a function of a large number of degrees of freedom, it is obviously impossible to keep track of all coordinates. Hence, to actually observe the folding funnel of the peptide, one has to study a projection of the landscape onto a set of suitable and appropriate order parameters. For our choice of the order parameters, we were guided by previous numerical simulations of Met-enkephalin.[29,34–41] Using the ECEPP/2 force field,[42] it was shown in a recent article[29] that Met-enkephalin undergoes a transition between extended and compact structures at a temperature $T_\theta = 295 \pm 20$ K. Above that temperature, the frequency of compact structures rapidly decreases while it increases below T_θ. Hence, our first order parameter is the volume allowing us to distinguish between compact and extended conformations. In ref. 29 it was also shown that by further lowering the temperature, the peptide encounters a second transition. Below $T_f = 230 \pm 30$ K, the occupation of the ground-state conformations increases rapidly while it decreases for values of T above T_f. The ensemble of low-temperature conformations was studied by various groups[23,34–37,39] and the numerical results compared with the experimental findings.[38] These studies agree in that there are two major groups of well-defined compact structures which are characterized (and stabilized) by specific hydrogen bonding patterns. In Figure 1 we show a sketch of the two structures. Structure A is the ground-state conformation in ECEEP/2 and has a Type II′ β-turn between the second and last residue, stabilized by two possible hydrogen bonds. The structure B, the second-lowest energy state, is characterized by hydrogen bond between Tyr-1 and Phe-4 resulting in a Type-II β-turn between the first and fourth residue. We remark that at higher energies (in ECEPP more than 2 kcal/mol above the ground state) conformations with a γ-turn and those with hydrogen bonding between the first and last residue were also observed.[36,41] Hence, we choose as further order parameters the overlap with the ground state (structure A) and the second-lowest-energy state (structure B), respectively, which allows us to distinguish between the various compact low-energy conformations.

In the following sections we first review our simulation method and the ECEPP force field and show afterwards how the above order parameters can be defined and measured. Our results are then presented and we finish with our conclusions.

SIMULATION TECHNIQUES

The generalized-ensemble technique utilized in this article was first introduced in refs. 43, 44. In this algorithm, configurations are updated according to the following probability weight:

$$w(E) = \left(1 + \frac{\beta(E - E_0)}{n_F}\right)^{-n_F}, \qquad (1)$$

where E_0 is an estimator for the ground-state energy, n_F is the number of degrees of freedom of the system, and $\beta = 1/k_B T$ is the inverse temperature (k_B is the Boltzmann constant and T the temperature of the system). Note that this weight is a special case of the weights used in Tsallis generalized mechanics formalism[45] (the Tsallis parameter q is chosen as $q = 1 + 1/n_F$). In the low-energy region (where $\beta(E - E_0)/n_F \ll 1$), the new weight reduces to the canonical Boltzmann weight $\exp(-\beta E)$. On the other hand, high-energy regions are no longer exponentially suppressed but only according to a power law, which enhances excursions to high-energy regions. We remark that the calculation of the weight is much easier than in other generalized-ensemble techniques, since it requires one to find only an estimator for the ground-state energy E_0, which can be done by a procedure described in ref. 44.

As in the case of other generalized ensembles, we can use the reweighting techniques[46] to construct canonical distribu-

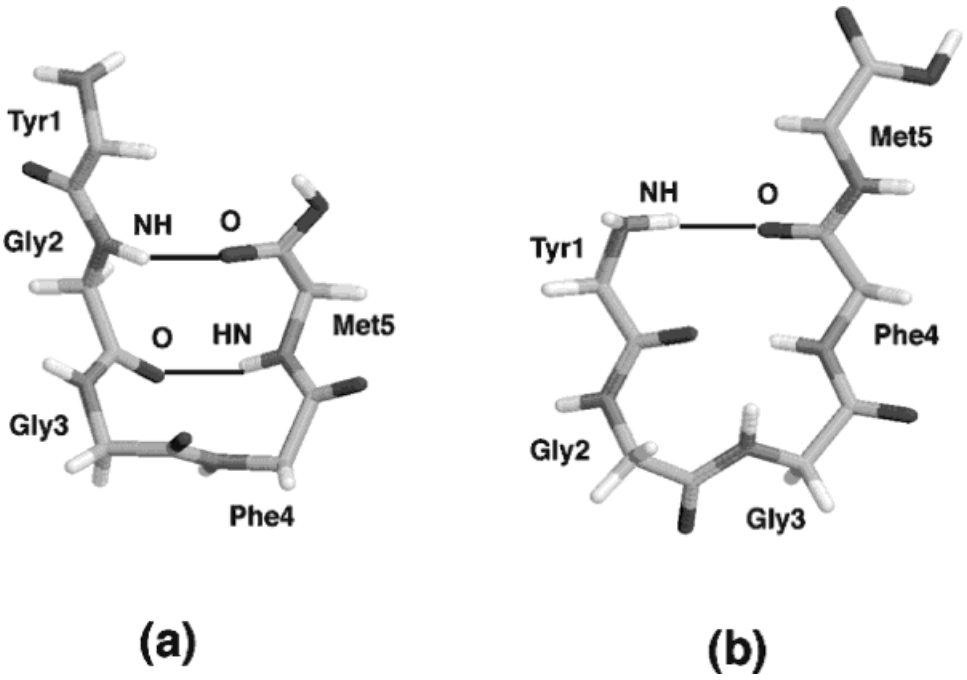

Fig. 1. Backbone structures of the two dominant low-energy structures with their characteristic hydrogen bonding. The figures were created with RasMol.[55] Part **a** displays conformer A, the ground state in ECEPP/2 (with a KONF90-energy of −12.2 kcal/mol). Conformer B in Part **b** is the local minimum with the second lowest potential energy (−11.0 kcal/mol in KONF90).

tions at various temperatures. This is because the simulation by the present algorithm samples a large range of energies. The thermodynamic average of any physical quantity $\mathcal{A}$ can be calculated over a wide temperature range by

$$\langle \mathcal{A} \rangle_T = \frac{\int dx\, \mathcal{A}(x) w^{-1}(E(x)) e^{-\beta E(x)}}{\int dx w^{-1}(E(x)) e^{-\beta E(x)}}, \tag{2}$$

where x stands for configurations.

FORCE FIELDS

For our simulations we used the ECEEP/2 force field[42] in which the potential energy function E_{tot} is given by the sum of the electrostatic term E_C, 12-6 Lennard-Jones term E_{LJ}, and hydrogen-bond term E_{HB} for all pairs of atoms in the peptide together with the torsion term E_{tor} for all torsion angles:

$$E_{tot} = E_C + E_{LJ} + E_{HB} + E_{tor}, \tag{3}$$

$$E_C = \sum_{(i,j)} \frac{332 q_i q_j}{\epsilon r_{ij}}, \tag{4}$$

$$E_{LJ} = \sum_{(i,j)} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right), \tag{5}$$

$$E_{HB} = \sum_{(i,j)} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right), \tag{6}$$

$$E_{tor} = \sum_{l} U_l (1 \pm \cos(n_l \chi_l)). \tag{7}$$

Here, r_{ij} (in Å) is the distance between the atoms i and j, and χ_l is the torsion angle for the chemical bond l. Bond lengths and bond angles are fixed at experimental values, leaving the dihedral angles ϕ, ψ, ω, and χ as independent variables. We further fix the peptide bond angles ω to their common value 180°, which leaves us with 19 torsion angles (ϕ, ψ, and χ) as independent degrees of freedom (i.e., $n_F = 19$). In our simulations we did not explicitly include the interaction of the peptide with the solvent and set the dielectric constant ϵ equal to 2. However, we do expect some implicit solvent effect, since the various parameters (q_i, A_{ij}, B_{ij}, C_{ij}, D_{ij}, U_l, and n_l) for the energy function were determined by minimization of the potential energies of the crystal lattices of single amino acids, i.e., not in a vacuum. We remark that the computer code KONF90[47] which we used in our simulation relies on a different convention for the implementation of the ECEPP parameters (for example, ϕ_1 of ECEPP/2 is equal to $\phi_1 - 180°$ of KONF90). Therefore, our energy values are slightly different from those of the original implementation of ECEPP/2.

TECHNICAL DETAILS

It is known from our previous work that the ground-state conformation for Met-enkephalin has the KONF90 energy value $E_{GS} = -12.2$ kcal/mol.[37] We therefore set $E_0 = -12.2$ kcal/mol, $T = 50$ K (or, $\beta = 10.1$ [1/kcal/mol]) and $n_F = 19$ in our probability weight factor in Eq. (1). Our simulation was started from a completely random initial conformation (Hot Start) and one Monte Carlo sweep updates every torsion angle of the peptide once. All thermodynamic quantities were then calculated from a single

production run of 1,000,000 MC sweeps which followed 10,000 sweeps for thermalization. At the end of every fourth sweep we stored the energies of the conformation and our three "order parameters" (the corresponding volume, the overlap O_A of the conformation with the (known) ground state (structure A) and the overlap O_B of the conformation with conformer B. Here, we approximate the volume of the peptide by its solvent excluded volume (in Å^3) which is calculated by a variant[48] of the double cubic lattice method.[49] Our definition of the overlap, which measures how much a given conformation resembles a reference state, is given by

$$O(t) = 1 - \frac{1}{90\ n_F} \sum_{i=1}^{n_F} |\alpha_i^{(t)} - \alpha_i^{(RS)}|, \tag{8}$$

where $\alpha_i^{(t)}$ and $\alpha_i^{(RS)}$ (in degrees) stand for the n_F dihedral angles of the conformation at t-th Monte Carlo sweep and the reference state conformation, respectively. Symmetries for the side-chain angles were taken into account and the difference $\alpha_i^{(t)} - \alpha_i^{(RS)}$ was always projected into the interval [−180°, 180°]. Our definition guarantees that we have

$$0 \leq \langle O \rangle_T \leq 1. \tag{9}$$

We remark that the average overlap $\langle O \rangle_T$ approaches its limiting value zero (for $T \to \infty$) only very slowly as the temperature increases. For instance, at $T = 1000$ K we found for the overlap with the ground state still has an average value of $\langle O_A \rangle \approx 0.3$. This is because $\langle O \rangle_T = 0$ corresponds to a completely random distribution of dihedral angles which is energetically highly unfavorable due to the steric hindrance of both main and side chains. Note the obvious limit: $O_A \to 1$, as $T \to 0$.

Since large parts of the configuration space are sampled by our method, the use of the reweighting techniques[46] is justified to calculate thermodynamic quantities over a wide range of temperatures by Eq. (2). Examples are the average potential energy $\langle E \rangle (V)$, the entropy

$$S(V) = \langle E \rangle (V)/k_B T + \log P(V) \tag{10}$$

(where $P(V)$ is the probability to find a conformation with the volume V), and similar quantities defined as functions of the two overlaps O_A and O_B. We normalized all the above quantities in such a way that they are zero for the ground state.

The above defined quantities allow only to study a projection of the energy landscape into one dimension. To get a more detailed picture, we also explored for various temperatures the free energies

$$G(O_A, O_B) = -k_B T \log P(O_A, O_B) \tag{11}$$

and

$$G(O_A, V) = -k_B T \log P(O_A, V), \tag{12}$$

where again $P(O_A, O_B)$ and $P(O_A, V)$ are respectively the probabilities to find a peptide conformation with values O_A, O_B and O_A, V. We chose the normalization so that the lowest value of $G(O_A, O_B)$ or $G(O_A, V)$ is set to zero for each temperature.

Finally, in order to monitor the number of states in which the peptide can be found at temperature T, we also calculated the (unnormalized) entropy

$$S(T) = \langle E \rangle (T)/k_B T - \log Z(T), \tag{13}$$

where $Z(T)$ is the estimate of the partition function of the system as calculated from our data by the reweighting techniques. This quantity allows one to estimate the glass temperature T_g, since the number of possible states should decrease drastically after entering the glassy phase.

THE MET-ENKEPHALIN FUNNEL LANDSCAPE

The energy landscape for a folding protein strongly depends on temperature. For some temperatures nonspecific trapping may not be a problem but the situation may completely reverse for other temperatures. Therefore, when exploring the folding landscape for this peptide, the simulations should take place at relevant temperatures, for otherwise one is not able to obtain a reliable picture of the folding mechanism. Hence, we have concentrated our analyses on four temperatures. The first one, $T = 1{,}000$ K was chosen to probe the high-temperature regime where the peptide is fully unfolded and mostly in an extended form. In some early work,[29] some of us have identified $T = 300$ K as the collapse temperature T_θ and $T = 230$ K as the folding temperature T_f. The last temperature, $T = 150$ K, was chosen to study the low temperature behavior of the peptide where the glassy behavior is observed.

Figure 2a displays the average potential energy $\langle E \rangle (V)$ as a function of the volume of the peptide for the four chosen temperatures. The plot shows that configurations with small volume (<1400 Å^3) have essentially the same energy. Above that value of the volume the energy increases with the volume. The increase is only gradual for high temperatures, but very steep below T_θ. Similar results were found for the entropy $S(V)$, displayed for the same four temperatures in Figure 2b. The number of states varies slightly below a certain threshold but increases rapidly above that value. The steepness of that increase is again a function of temperature below T_θ. The two plots indicate that one can distinguish between two regimes: compact structures which have similar energies and entropy and extended structures which are entropically favored but energetically disfavored.

Figure 3a (3b) shows the average energy (entropy) as a function of the overlap with structure A (the ground state in ECEPP/2). For the interpretation of the plots one has to keep in mind that the overlaps approach zero very slowly with increasing energy or temperature. Results from our simulation are only trustworthy for temperatures $T \leq 1{,}000$ K where the overlap function has average value of ≈0.3. Hence, in Figures 3a and 3b the results are only reliable for values of the overlap function above 0.3.

At high temperatures, the dependence of $\langle E \rangle (O_A)$ on the overlap O_A is essentially a monotonous function where the energy decreases with increasing value of the overlap. Below the collapse temperature T_θ, however, the curves show a minimum at $O_A \approx 0.5$. This is the value O_A of

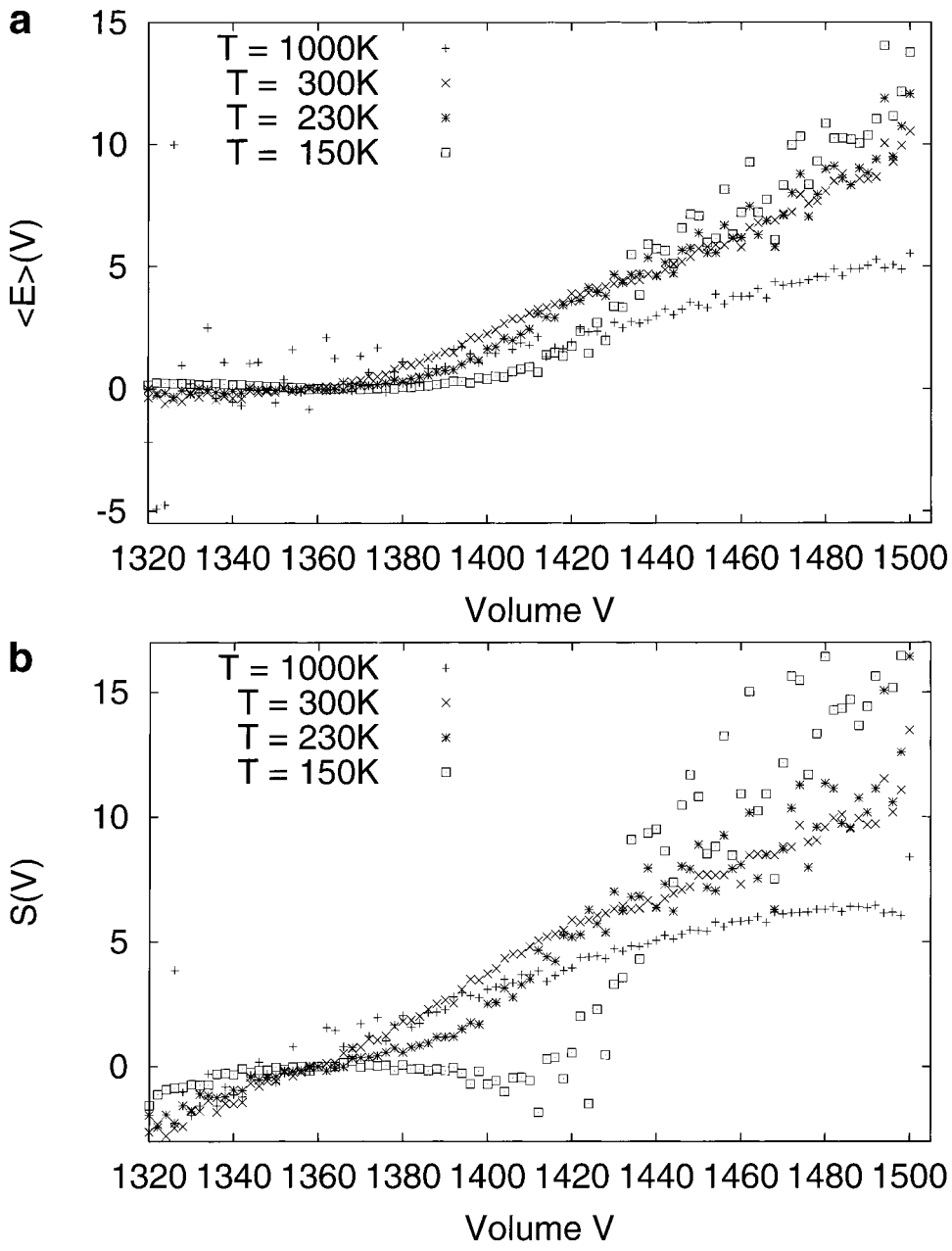

Fig. 2. (**a**) Potential energy $\langle E \rangle (V)$ (in kcal/mol) and (**b**) entropy $S(V)$ (as defined in Eq. (10)) as a function of the volume (in Å^3) of the peptide for various temperatures (in K). We chose units where for the ground state $\langle E \rangle (V) = 0$ and $S(V) = 0$.

Conformation B (the exact value being 0.46), and this minimum is therefore an indication of a basin of attraction to Conformer B. However, the overall slope of $\langle E \rangle$ (O_A) shows that Conformer A is favored. On the other hand, the entropy as a function of the overlap strongly varies with decreasing overlap for all but the lowest temperature. The absolute values of the entropy decrease with temperature. At $T = 150$ K, however, $S(O_A)$ stays essentially constant over the whole range of values of the overlap once this value is only a little different from 1. This means that at this temperature the number of states differs little with the overlap and therefore indicates the onset of glassy behavior. Similar pictures hold when one prints energy and entropy differences as function of the overlap with Conformation B. This kind of behavior has been predicted from simulations of minimalist models and now they have been confirmed for simulations of a real peptide.

As the number of available states gets reduced with the decrease of temperature, the possibility of local trapping increases substantially. In the thermodynamic limit, the system would be trapped in these local traps for an infinite time as expected for a glass transition. To determine the

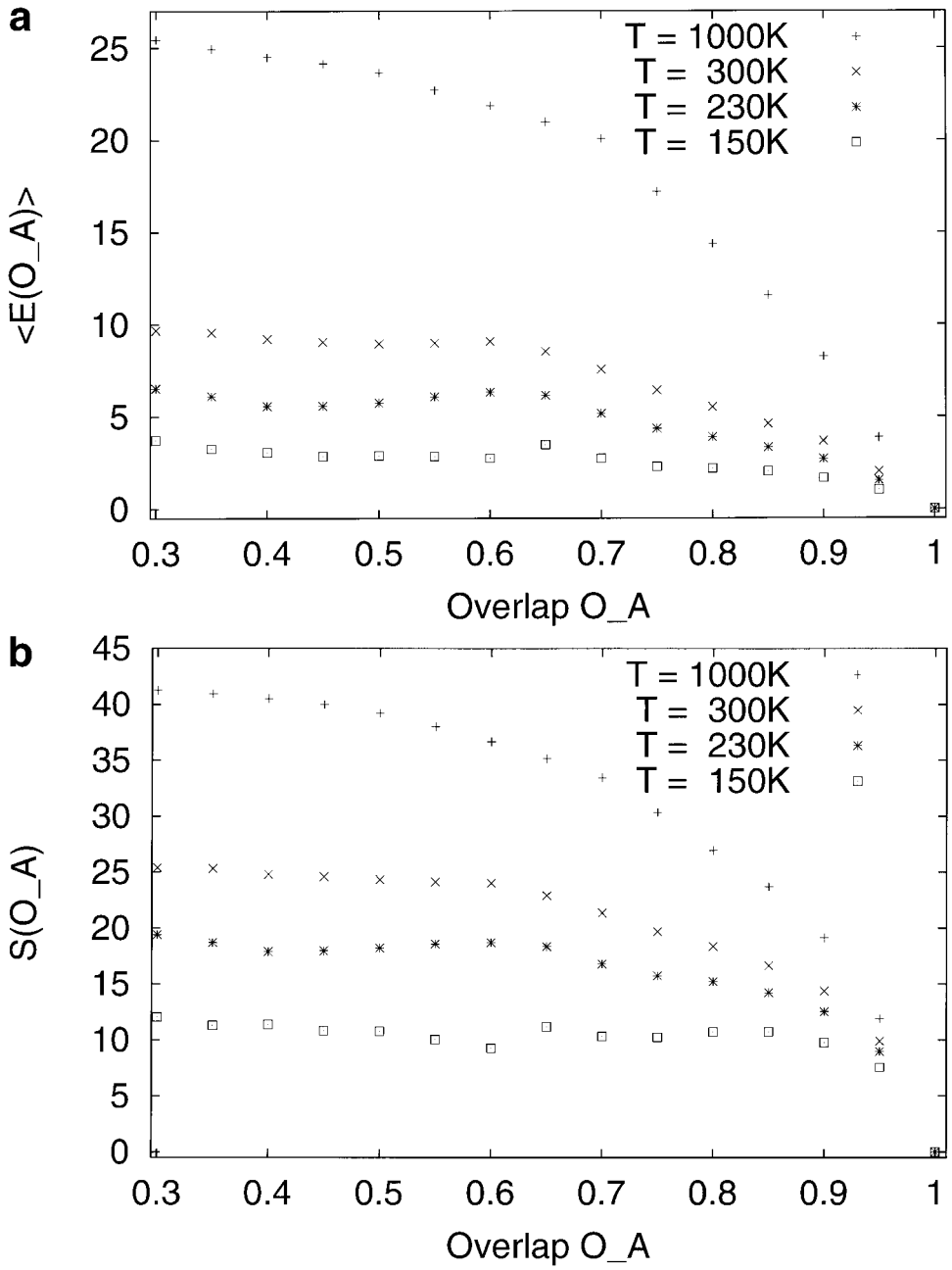

Fig. 3. (**a**) Potential energy $\langle E \rangle (O_A)$ (in kcal/mol) and (**b**) entropy $S(O_A)$ as a function of the overlap O_A (defined in the text) for various temperatures (in K). We chose units where for the ground state $\langle E \rangle (O_A) = 0$ and $S(O_A) = 0$.

glassy behavior for a finite system is a more subtle problem. In the folding problem, glassy behavior is associated when the residence time in some local traps becomes of the order of the folding event. Folding dynamics is now nonexponential since different traps have different escape times.[15,50] For temperatures above the glass transition temperature T_g, the folding dynamics is exponential and a configurational diffusion coefficient average the effects of the short lived traps.[10,51,52] It is expected that for a good folder the temperature, T_g, where glass behavior sets in, has to be significantly lower than the folding temperature T_f, i.e. a good folder can be characterized by the relation[5]

$$\frac{T_f}{T_g} > 1. \tag{14}$$

Our plots of $\Delta S(O_A)$ indicate that we indeed have $T_g < T_f$. This is also supported by Figure 4 where the (unnormalized) entropy of the molecule is represented as a function of

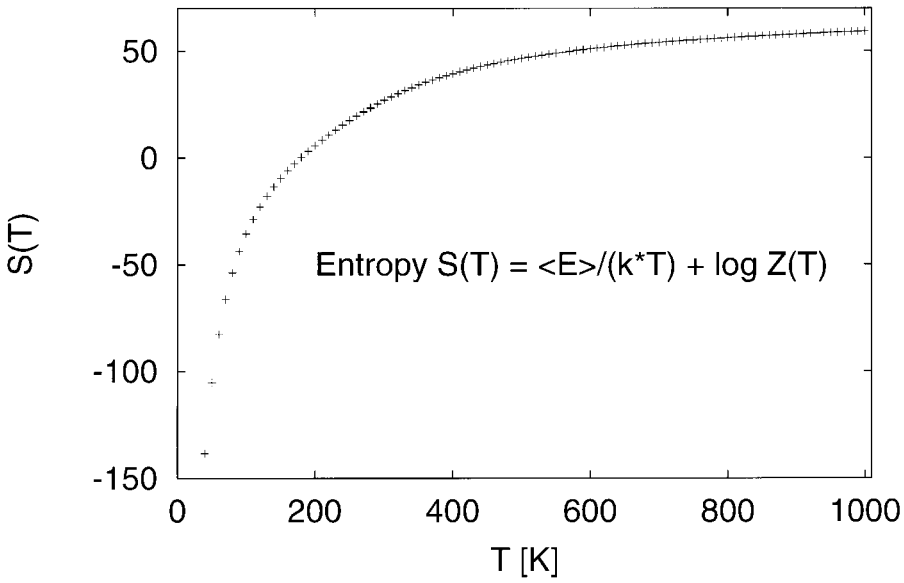

Fig. 4. Unnormalized entropy $S(T)$ (as defined in Eq. (13)) as a function of temperature (in K).

temperature. This plot clearly shows the rapid decrease of entropy once T reaches values smaller than T_g. As one would expect from such a small molecule, it is impossible to determine a precise value for T_g from the plot, but clearly the onset of glassy behavior happens at temperatures well below the folding temperature $T_f = 230$ K. In ref. 29 it was already pointed out that these results are also consistent with an alternative characterization of folding properties. Thirumalai and collaborators[53,54] have conjectured that the kinetic accessibility of the native conformation can be classified by the parameter

$$\sigma = \frac{T_\theta - T_f}{T_\theta}, \tag{15}$$

i.e., the smaller σ is, the more easily a protein can fold. With our values for T_θ and T_f, we have for Met-enkephalin $\sigma \approx 0.2$. Here, we have taken the central values: $T_\theta = 295$ K and $T_f = 230$ K. This value of σ implies that our peptide has reasonably good folding properties according to refs. 53 and 54. Hence, we see that there is a strong correlation between the folding criterion ($T_f/T_g > 1$) proposed by Bryngelson and Wolynes[5] and this one by Thirumalai and collaborators.[53,54]

The main goal of this article is to depict the folding funnel of Met-enkephalin. Since it is not feasible to plot the free energy G as a function of all three order parameters, one has to plot G as a function of a suitable combination of the three relevant order parameters of the molecule. We chose to plot the free energy $G(V, O_A)$ as a function of volume V and overlap O_A with the known ground state (Conformer A) and $G(O_A, O_B)$ as a function of the overlap with the ground state (O_A) and with Conformer B (O_B). Again we study these quantities for temperatures $T = 1{,}000$ K, $T = T_\theta = 300$ K, $T = T_f = 230$ K, and $T = 150$ K. They are shown in Figure 5a–d and Figure 6a–d, respectively.

In Figure 5a we show the free energy landscape as a function of volume and overlap with the known ground state (structure A) at the high-temperature region ($T = 1{,}000$ K). Here, (as in the other free energy plots) we normalized the free energy in such a way that its observed minimum is set to zero. In the contour plots, the contour lines mark multiples of k_BT (therefore different for different temperatures but appropriate to understand the folding mechanism). We see that the free energy has its minimum at large volumes ($\approx$1470 Å^3) and values of the overlap $O_A \approx 0.3$. Small volumes and larger values of the overlap are suppressed by many orders of k_BT. Hence, extended random coil structures are favored at this temperature. The picture changes dramatically once we reach the collapse temperature T_θ, shown in Figure 5b. At this temperature a large part of the V-O_A space can be sampled in a simulation. The contour plot shows that regions with both small and large volumes and almost all values of O_A lie within the 2 k_BT contour. This indicates that at this temperature the cross over between extended and compact structures happens with a small thermodynamic barrier between them. By lowering the temperature to $T_f = 230$K (determined in ref. 29), we now observe strong evidence for a funnel-like landscape (Fig. 5c). At this temperature the drive towards the native configuration is dominant and no long-lived traps exist. There is clearly a gradient towards the ground-state structure ($O_A \approx 1$), but other structures with similar volume (characterized by values of $O_A \approx 0.5$) are only separated by free energy barriers of order 1 k_BT. Below this temperature we expect that the ground state is clearly favored thermodynamically and separated from other low energy states by free energy barriers of many orders of k_BT. This can be seen in Figure 5d where at $T =$ 150K where other low energy states have free energies of 3 k_BT higher than the ground state and are separated by an additional barrier of 2 k_BT.

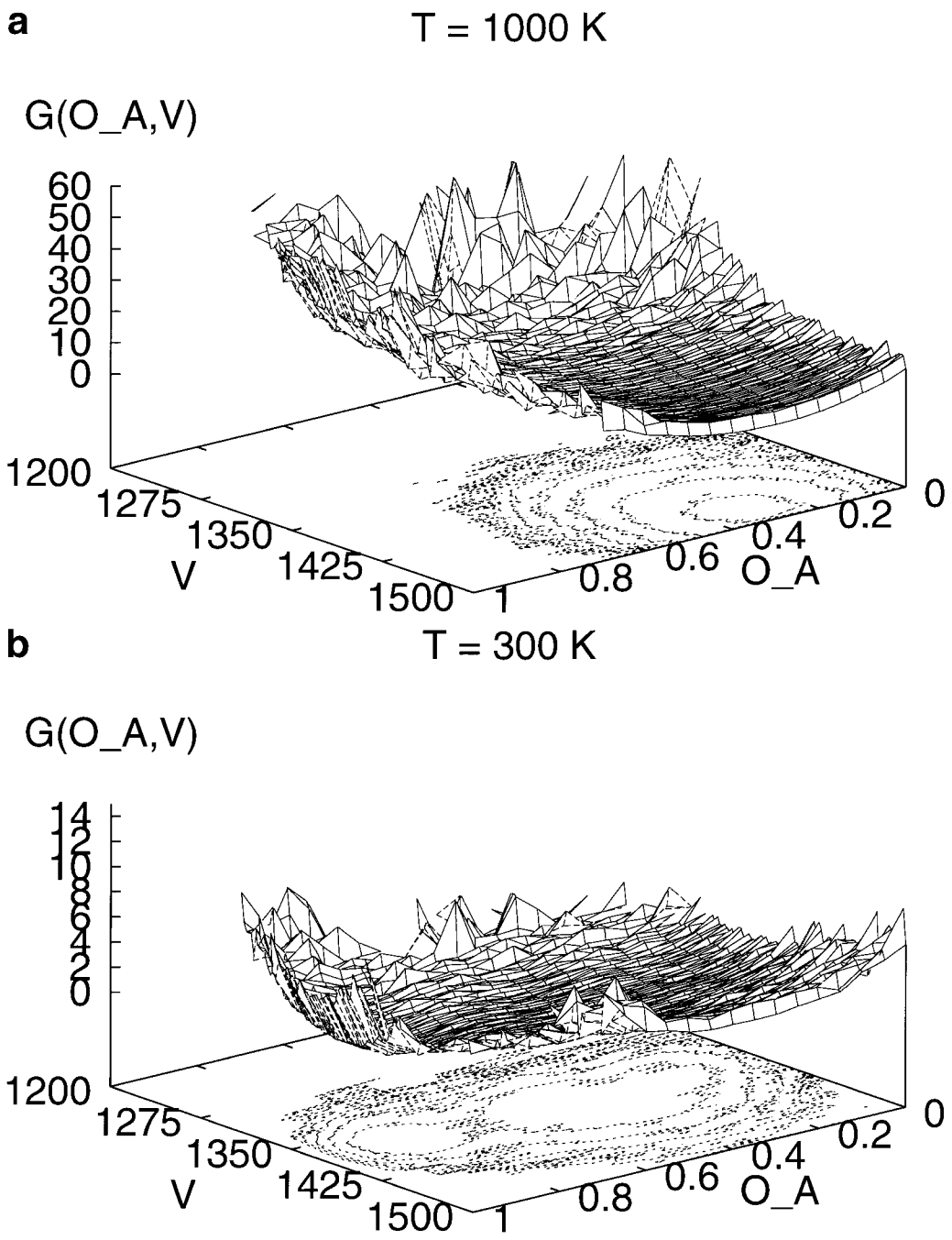

Fig. 5. Free energy $G(V, O_A)$ (in kcal/mol) as a function of both peptide volume V (in Å^3) and overlap O_A (as defined in the text) for (**a**) $T = 1000$ K, (**b**) $T = 300$ K, (**c**) $T = 230$ K, and (**d**) $T = 150$ K. Both the free energy surface and the contour plot are shown. The contour lines are multiples of k_BT. $G(V, O_A)$ was normalized such that $\min(G(V, O_A)) = 0$.

The above picture is supported by the plots for the free energy as a function of both the overlap O_A with the ground state and the overlap O_B with structure B. Figure 6a shows again the high-temperature situation. The free energy has its minimum at small values of the overlap indicating that both conformers appear with only very small frequency at high temperature. At $T = 300$K, the collapse temperature, again a large part of the space of possible configurations (characterized by values of O_A and O_B) lies within the $2k_BT$ contour as is clear from Figure 6b. At the folding temperature $T_f = 230$ K a funnel in the energy landscape appears with a gradient towards the ground state, but Figure 6c shows that there are various other structures, the most notable of which is Conformer B (where $O_B \approx 1$), with free

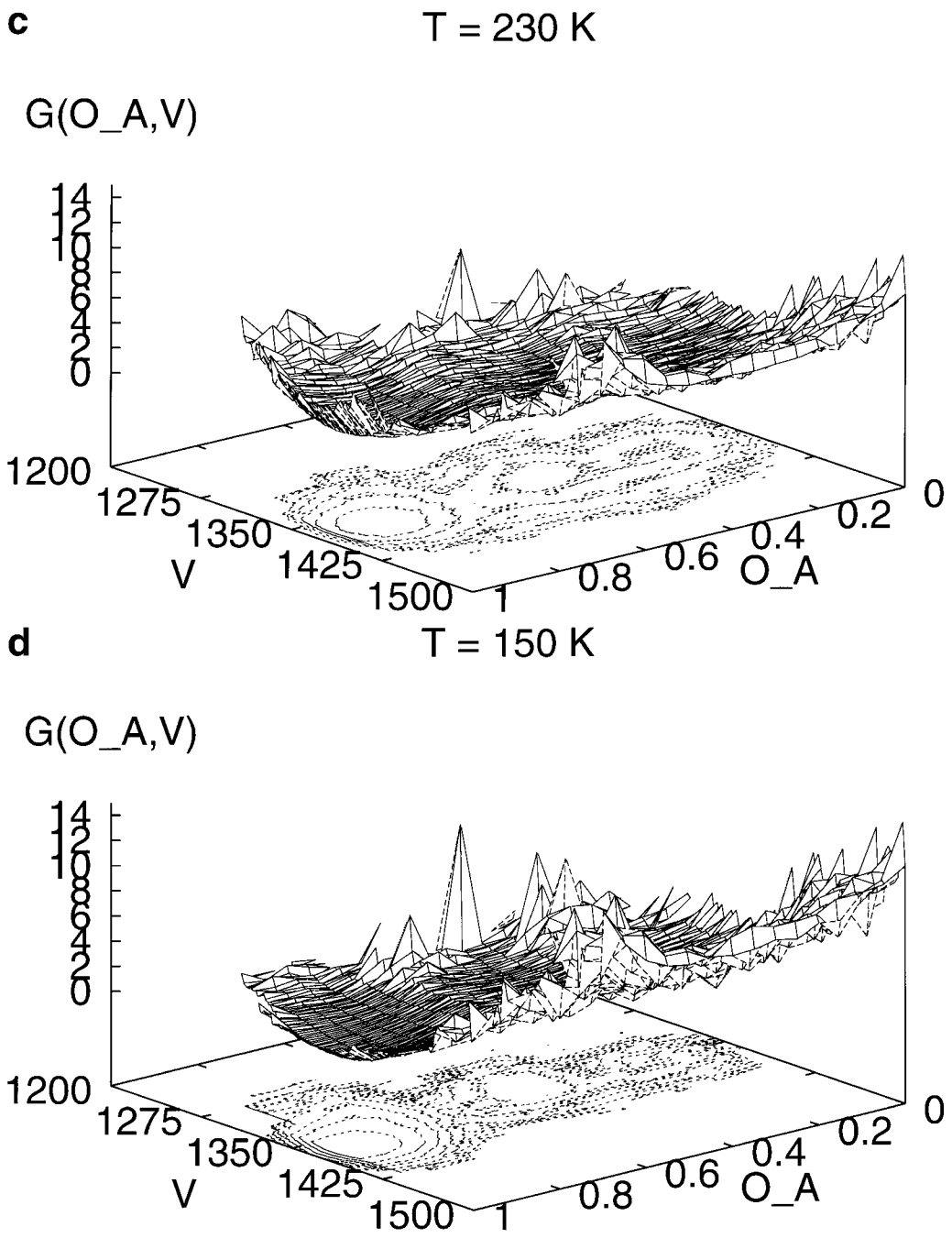

Figure 5. (Continued.)

energies 3 k_BT higher than the ground-state conformation but separated from each other and the ground state only by free energy barriers less than 1 k_BT. No other long-lived traps are populated. Hence, the funnel at T_f is reasonably smooth. Folding routes include direct conversion from random-coil conformations into Conformer A or some short trapping in Conformer B region before reaching Conformer A region, but at the folding temperature it is possible to reach the ground state from any configuration without getting kinetically trapped. Kinetic Monte Carlo runs at a fixed temperature ($T = 230$ K) were performed and confirmed this picture (data not shown). We observed that some of the runs went directly from the unfolded state to the ensemble of folded conformations in state A, while in other runs trapping at state B occurred first before folding into the ground-state structure. The kinetic runs therefore support our observation that Met-enkephalin is a good folder and that $T_f > T_g$. Figure 6d shows the situation for $T = 150$ K where we expect onset of glassy behavior. Again one sees a funnel-like bias toward the ground state,

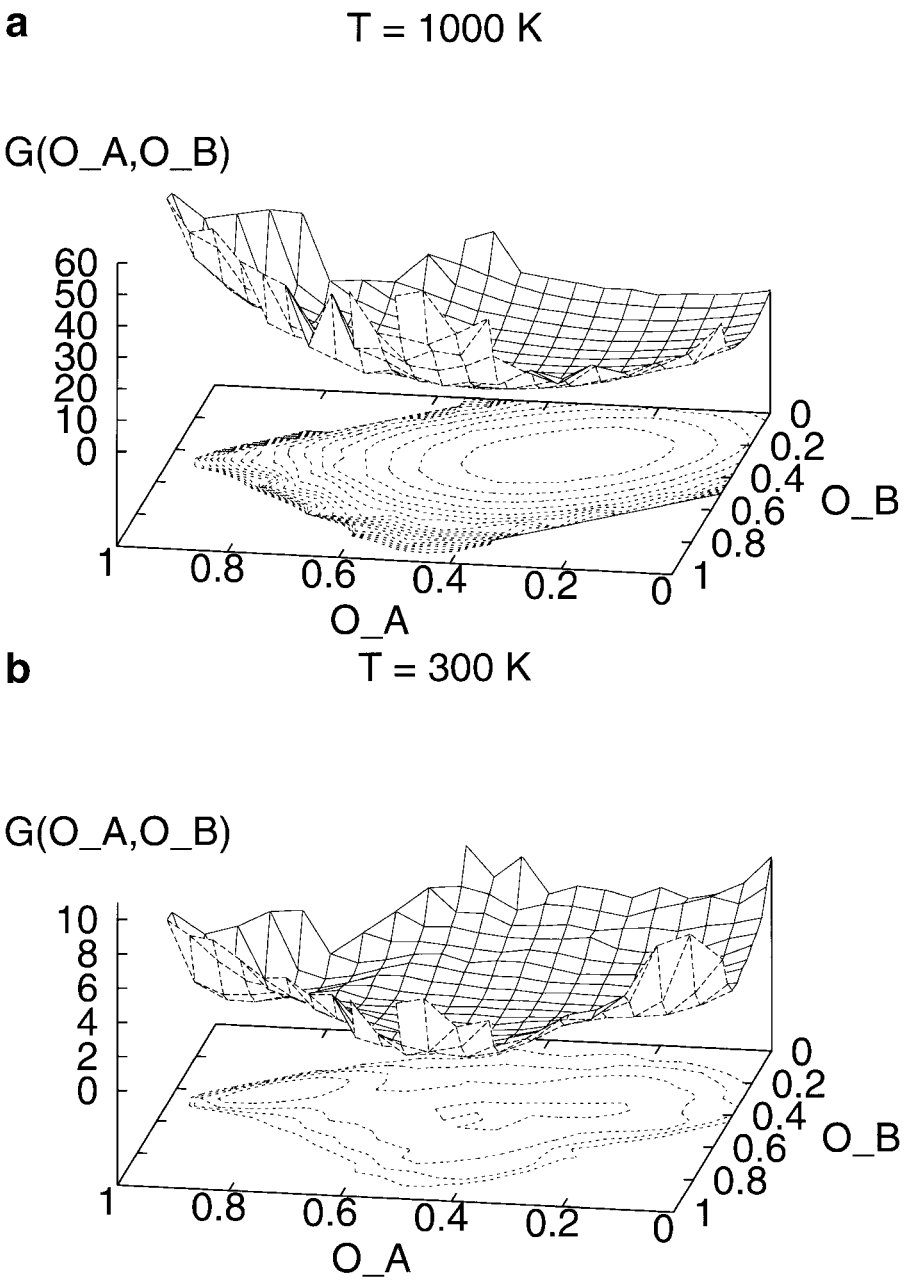

Fig. 6. Free energy $G(O_A, O_B)$ as a function of both overlaps O_A and O_B (as defined in the text) for (**a**) $T = 1000$ K, (**b**) $T = 300$ K, (**c**) $T = 230$ K, and (**d**) $T = 150$ K. Both the free energy surface and the contour plot are shown. The contour lines are multiples of k_BT. $G(O_A, O_B)$ was normalized such that $\min(G(O_A, O_B)) = 0$.

however, the funnel is no longer smooth and the free energy landscape is rugged. Free energy barriers of many k_BT now separate different regions and would act as long-lived kinetic traps in a canonical simulation, rendering folding at this temperature extremely difficult.

CONCLUSIONS

In this article we have studied the free energy landscape of the peptide Met-enkephalin, using generalized-ensemble techniques. Although the peptide is rather small, the obtained free energy landscape shows a funnel towards the ground state even for temperatures well below the folding temperature T_f. It was shown that glassy behavior appears only well below this temperature and that the peptide is a good folder at T_f. Our results demonstrate that the energy landscape picture and funnel concept of folding that have been demonstrated in the past for minimalist models are much more general and hold for this peptide when simulated with an all-atom force field.

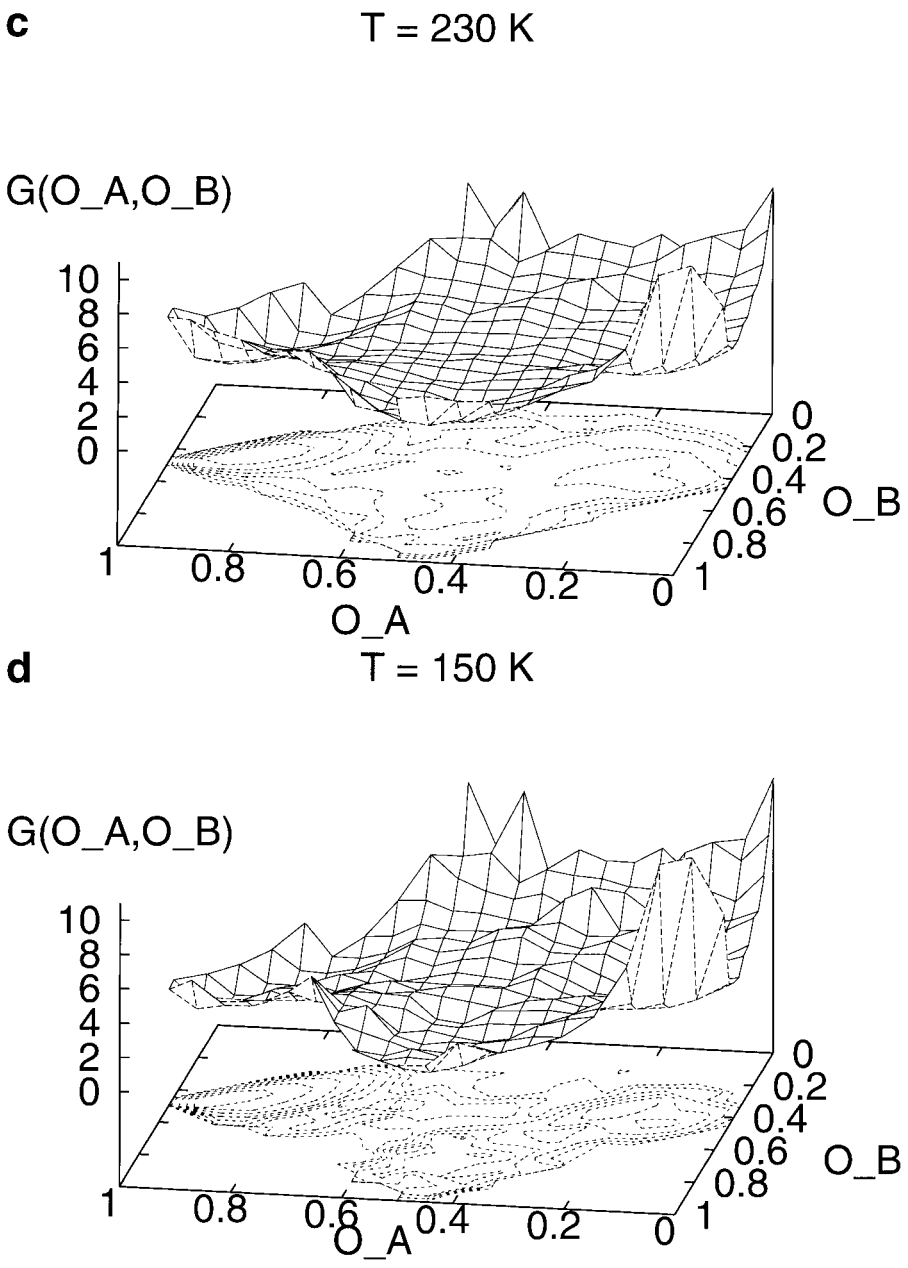

Figure 6. (Continued.)

ACKNOWLEDGMENTS

Our simulations were performed on computers of the Institute for Molecular Science (IMS), Okazaki, Japan. This work is supported by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Science, Sports and Culture. Work in San Diego was also supported by the National Science Foundation (Grant number 96-03839). Work in Houghton was supported by a Research Excellence Fund (E27448) of the State of Michigan. JNO was a visitor at the IMS when part of this work was performed and he thanks Y. Tanimura for his hospitality during his stay in Japan.

REFERENCES

1. Anfinsen CB. Principles that govern the folding of protein chains. Science 1973;181:223–230.
2. Dill KA, Chan HS. From Levinthal to pathways to funnels. Nat Struct Biol 1997;4:10–19.
3. Sali A, Shakhnovich EI, Karplus M. Kinetics of protein folding—A lattice model study of the requirements for folding to the native state. J Mol Biol 1994;235:1614–1636.
4. Guo Zh, Thirumalai D. Kinetics and thermodynamics of folding of a de novo designed four-helix bundle protein. J Mol Biol 1996;236: 323–343.
5. Bryngelson JD, Wolynes PG. Spin glasses and the statistical mechanics of protein folding. Proc Natl Acad Sci USA 1987;84: 7524–7528.
6. Leopold PE, Montal M, Onuchic JN. Protein folding funnels:

kinetic pathways through compact conformational space. Proc Natl Acad Sci USA 1992;89:8721–8725.
7. Onuchic JN, Luhey-Schulten Z, Wolynes PG. Theory of protein folding: the energy landscape perspective. Annu Rev Phys Chem 1997;48:545–600.
8. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways and the energy landscape of protein folding: a synthesis. Proteins 1995;21:167–195.
9. Onuchic JN, Wolynes PG, Luthey-Schulten Z, Socci ND. Towards an outline of the topography of a realistic protein folding funnel. Proc Natl Acad Sci USA 1995;92:3626–3630.
10. Socci ND, Onuchic JN, Wolynes PG. Diffusive dynamics of the reaction coordinate for protein folding funnels. J Chem Phys 1996;104:5860–5871.
11. Guo Zh, Thirumalai D. Kinetics of protein folding—nucleation mechanism, time scales, and pathways. Biopolymers 36:103–116, 1995.
12. Mirny LA, Abkevich V, Shakhnovich EI. Universality and diversity of the protein folding scenarios: a comprehensive analysis with the aid of a lattice model. Fold Des 1996;1:103–116.
13. Friedrichs MS, Goldstein RA, Wolynes PG. Generalized protein tertiary structure recognition using associative memory hamiltonians. J Mol Biol 1991;222:1013–1034.
14. Guo Zh, Brooks III, CL. Thermodynamics of protein folding: a statistical mechanical study of a small all β protein. Biopolymers 1997;42:745–757.
15. Nymeyer H, García AE, Onuchic JN. Folding funnels and frustration in off-lattice minimalist protein landscapes. Proc Natl Acad Sci USA 1998;95:5921–5928.
16. Shea J-E, Nochomovitz YD, Guo Zh, Brooks III, CL. Exploring the space of protein folding hamiltonians: the balance of forces in a minimalist β-barrel model. J Chem Phys 1998;109:2895-2903.
17. Hardin C, Luthey-Schulten ZA, Wolynes PG. Backbone dynamics, fast folding, and secondary structure of helical proteins and peptides. Proteins 1998; in press.
18. Wolynes PG, Luthey-Schulten ZA, Onuchic JN. Fast-folding experiments and the topography of protein folding energy landscapes. Chem Biol 1996;3:425–432.
19. Vásquez M, Némethy G, Scheraga HA. Conformational energy calculations of polypeptides and proteins. Chem Rev 1994;94:2183–2239.
20. Hansmann UHE, Okamoto Y. The generalized-ensemble approach for protein folding simulations. In: Stauffer D, editor. Annual reviews in computational physics VI. Singapore: World Scientific; 1998, in press.
21. Berg BA, Neuhaus T. Multicanonical algorithms for first order phase transitions. Phys Lett 1991;267:249–253.
22. Lyubartsev AP, Martinovski AA, Shevkunov SV, Vorontsov-Velyaminov PN. New approach to Monte Carlo calculations of the free energy: method of expanded ensembles. J Chem Phys 1992;96: 1776–1783; Marinari E, Parisi G. Simulated tempering: a new Monte Carlo scheme. Europhys Lett 1992;19:451-458.
23. Hansmann UHE, Okamoto Y. Prediction of peptide conformation by multicanonical algorithm: a new approach to the multiple-minima problem. J Comp Chem 1993;14:1333–1338.
24. Okamoto Y, Hansmann UHE. Thermodynamics of helix—coil transitions studied by multicanonical algorithms. J Phys Chem 1995;99:11276–11287.
25. Hansmann UHE, Okamoto Y. Finite-size scaling of helix-coil transitions in poly-alanine studied by multicanonical simulations. J Chem Phys 1999; in press.
26. Hansmann UHE, Okamoto Y. Tertiary structure prediction of C-peptide of ribonuclease A by multicanonical algorithm. J Phys Chem 1998;102:653–656.
27. Hansmann UHE, Okamoto Y. Effects of side-chain charges on α-helix stability in C-peptide of ribonuclease A studied by multicanonical algorithm. J Phys Chem 1999; in press.
28. Hansmann UHE, Okamoto Y. Numerical comparisons of three recently proposed algorithms in the protein folding problem. J Comp Chem 1997;18:920–933.
29. Hansmann UHE, Masuya M, Okamoto Y. Characteristic temperatures of folding of a small peptide. Proc Natl Acad Sci USA 1997;94:10652–10656.
30. Boczko EM, Brooks CL III. First principles calculation of the folding free energy for a three helix bundle protein. Science 1995;269:393–396.
31. Aubry A, Birlirakis N, Sakarellos-Daitsiotis M, Sakarellos C. A crystal molecular conformation of leu-enkephalin related to the morphin molecule. Biopolymers 1989;28:27–40.
32. Behnam BA, Deber CM. Evidence for a folded conformation of methionine- and Leucine-Enkephalin in a Membrane Environment. J Biol Chem 1984;259:14939–14940.
33. Karayannis T, Gerothanassis IP, Sakarellos-Daitsiotis M, Sakarellos C, Marraud M. ^{17}O and ^{14}N-NMR studies of leu-enkephalin and enkephalin related fragments in aqueous solution. Biopolymers 1990;29:423–439.
34. Li Z, Scheraga HA. Monte Carlo minimization approach to the multiple-minima problem in protein folding. Proc Natl Acad Sci USA 1984;84:6611–6615.
35. Freyberg B von, Braun W. Efficient search for all low energy conformations of polypeptides by Monte Carlo methods. J Comp Chem 1991;12:1065–1076.
36. Okamoto Y, Kikuchi T, Kawai H. Prediction of low-energy structures of met-enkephalin by Monte Carlo simulated annealing. Chem Lett 1992;1992:1275–1278.
37. Hansmann UHE, Okamoto Y. Comparative study of multicanonical and simulated annealing algorithms in the protein folding problem. Physica A 1994;212:415–437.
38. Montcalm T, Cui W, Zhao H, Guarnieri F, Wilson SR. Simulated annealing of met-enkephalin: low energy states and their relevance to membrane-bound, solution and solid-state conformations. J Mol Struct Theochem 1994;308:37–51.
39. Eisenmenger F, Hansmann UHE. Variation of the energy landscape of a small peptide under a change from the ECEPP/2 Force Field to ECEPP/3. J Phys Chem B 1997;101:3304–3310.
40. Hansmann UHE. Parallel tempering algorithm for conformational studies of biological molecules. Chem Phys Lett 1997;281: 140–150.
41. Mitsutake A, Hansmann UHE, Okamoto Y. in preparation.
42. Sippl MJ, Némethy G, Scheraga HA. Intermolecular potentials from crystal data. 6. Determination of empirical potentials for O-H $\cdots$ O = C hydrogen bonds from packing configurations. J Phys Chem 1994;88:6231–6233. (And references therein.)
43. Hansmann UHE. Simulated annealing with Tsallis weights—A numerical comparison. Physica A 1997;242:250–257.
44. Hansmann UHE, Okamoto Y. Generalized-ensemble Monte Carlo method for systems with rough energy landscape. Phys Rev E 1997;56:2228–2233.
45. Tsallis C. Possible generalization of Boltzmann-Gibbs statistics. J Stat Phys 1988;52:479–487.
46. Ferrenberg AM, Swendsen RH. Optimized Monte Carlo data analysis. Phys Rev Lett 1988;61:2635–2638. Phys Rev Lett 1989; 63:1658–1658(E). (And references given in the erratum.)
47. Kawai H, Okamoto Y, Fukugita M, Nakazawa T, Kikuchi T. Prediction of α-helix folding of isolated C-peptide of ribonuclease A by Monte Carlo simulated annealing. Chem Lett 1991;213–216.
48. The program for the calculation of solvent excluded volume was written by M. Masuya and will be described in detail elsewhere.
49. Eisenhaber F, Lijnzaad P, Argos P, Sander C, Scharf M. The double cubic lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. J Comp Chem 1995;16:273–284.
50. Socci ND, Onuchic JN, Wolynes PG. Protein folding mechanisms and the multidimensional folding funnel. Proteins 1998;32:136–158.
51. Scalley ML, Baker D. Protein folding kinetics exhibit an Arrhenius temperature dependence when corrected for the temperature dependence of protein stability. Proc Natl Acad Sci USA 1997;44: 10636–10640.
52. Plotkin SS, Wolynes PG. Non-Markovian configurational diffusion and reaction coordinates in protein folding. Phys Rev Lett 1998;22: 5015–5018.
53. Camacho CJ, Thirumalai D. A criterion that determines fast folding of proteins: A model study. Proc Natl Acad Sci USA 1993;90:6369–6372.
54. Klimov DK, Thirumalai D. Factors governing the foldability of proteins. Proteins 1996;26:411–441.
55. Sayle RA, Milner-White EJ. RasMol: biomolecular graphics for all. TIBS 1995;20:374–376.