ELSEVIER

Biophysical Chemistry 85 (2000) 93–100

Biophysical Chemistry

www.elsevier.nl/locate/bpc

Kinetics of a finite one-dimensional spin system as a model for protein folding

Takeshi Kikuchi*,[1]

Frontier Research Center for Computational Sciences, Science University of Tokyo, 2641 Yamazaki, Noda, Chiba 278-8510, Japan

Received 5 October 1999; accepted 24 March 2000

Abstract

Simple spin models are used to analyze the kinetic nature of lowest energy state formation of the spin systems as models of protein folding kinetics. The models employed in the present paper are based on the spin systems as models of biopolymers previously proposed by the author for the analysis of the equilibrium nature of transitions [T. Kikuchi, Biophys. Chem. 65 (1997) 109]. In particular, the effect of frustrations on the kinetics is investigated with the Monte Carlo simulations in this study. The results show that the kinetics of the present systems are characterized by the ratio of foldables (pathways on the energy landscape that to lead to the lowest energy state) and the temperature dependence of the mean first passage time of foldables. We also discuss the free energy profile of the present models and the relation of the present results to the kinetics of actual proteins.

Keywords: Protein folding kinetics; Spin model; Monte Carlo simulation; Lowest energy structure formation; Foldability

1. Introduction

A protein folds to its native structure within a rather short time compared to the estimation from simple consideration as described in Levinthal's paradox [1]. To understand the mechanism of the rapid protein folding is a substantial problem in molecular biophysics. The funnel model is an idea currently accepted for this problem [2,3]. That is, in the coarse-grained view of an energy landscape of a protein, the fast folding should be characterized by the folding funnel, on which strong forces drive the structure to the native conformation at a high rate. With this view of protein folding, we have the further question: how can some kinds of polymers such as proteins produce a smooth energy landscape with a funnel

* Tel.: +81-86-440-1090; fax: +81-86-440-1062.
E-mail address: tkikuchi@chem.kusa.ac.jp (T. Kikuchi)
[1] Present address: Department of Chemical Technology, College of Science and Industrial Technology, Kurashiki University of Science and the Arts, 2640 Nishinoura, Tsurajima-cho, Kurashiki, Okayama 712-8505, Japan.

0301-4622/00/$ - see front matter
PII: S0301-4622(00)00152-6

in spite of the high degree of conformational freedom? A clue to this question seems to be the minimal frustration principle initially proposed by Bryngelson and Wolynes [4] and Go [5], i.e. a protein with minimal frustration might possess a smooth downhill energy landscape. It is plausible that protein sequences are designed through the long evolutionary process to minimize their frustrations. Thus examination of the effects of frustrations is a significant problem for understanding the mechanism of rapid protein folding. The effect of frustrations on energy landscapes of proteins is mainly studied with mean field treatment of proteins [6,7], but it is rather difficult to estimate the magnitude of frustrations in an actual protein system, or even in a model protein. Hence, the folding funnel, the energy landscape and the minimal frustration principle are useful concepts in the problem of protein folding, but we are still far from the essential understanding of the protein folding mechanism.

Furthermore, it is significant to define 'foldability of a protein' in relation to the concepts mentioned above; in other words, to elucidate how proteins with high foldability are differentiated quantitatively from other random heteropolymers. Several criteria of foldability have been proposed in the literature [7–11]. Those ideas might be practical and useful to specify the foldability for each protein, but the physical basis of the foldability is not entirely clear. We think that the foldability should be understood and defined based on the concept of the energy landscape, frustrations and so on. More attention should be paid to distinguish between the kinetic and equilibrium properties. For example, it would be valuable to clarify how transition temperature, which is clearly an equilibrium parameter, can be related to folding rate (τ), i.e. the free energy barrier which is a characteristic parameter in kinetics. Several good correlations between these parameters have been published [9,10].

Recently, to analyze the effect of frustrations in detail, we proposed [12] a simple model to simulate basic equilibrium folding properties of biopolymers including proteins and nucleic acids. This model is constituted of a spin system with energy parameters adjusted in several ways. The energy parameters with no frustration give a model showing a two-state transition in the van't Hoff plot between the lowest energy state and disordered states as expected from the minimal frustration principle in protein folding, when the transition occurs between the randomly frozen states and averaged disordered states in a model system with randomly selected energy parameters. When the frustrations are added gradually to the no-frustration system, the transition property gradually loses the two-state nature and exhibits the random nature. This shift of the property of the transition is characterized by the relative values of inflection points of coarse-grained energy and entropy curves as functions of an order parameter specifying a spin state, θ_a, i.e. a system shows clearly two-state transition if $\theta_a|_{Er} < \theta_a|_{Sr}$, where $\theta_a|_{Er}$ and $\theta_a|_{Sr}$ are the values of inflection points of the energy and entropy curve, respectively. This gives a kind of a criterion of the 'foldability' of a spin system in equilibrium. Thus, by virtue of the simplicity of the model, we can handle the effect of frustrations on the order–disorder transition, and the results serve to shed light on the actual protein folding mechanism. While we focused on the equilibrium transition of the spin system in the previous paper, it is also interesting to investigate the effect of frustrations on the rate of formation of the lowest energy spin state of the system, which corresponds to the protein folding rate. Hence, the main topic of this paper is the analysis of kinetics of our spin system. In particular, we focus on the effects of frustrations on kinetic behavior of lowest energy structure formation which is a model of protein folding, and analyze the main difference between a 'foldable' system (a system attaining to the lowest energy state) and a random system. We also discuss the relationship of our results to the actual protein folding kinetics.

2. Method

2.1. Model

We use the same model spin system proposed in the previous work [12] for the present problem.

That is, N spins are connected as a one-dimensional chain. The Hamiltonian of this system is expressed in Eq. (1).

$$H = \sum_{i>j}^{N} a\eta_i\sigma_i + \sum bJ_{ij}\sigma_i\sigma_j + \sum_{\alpha=3}^{N-1}\sum_{i=1}^{N-\alpha+1} bK_i^{\alpha}\prod_{\kappa=i}^{i+\alpha-1}\sigma_\kappa \quad (1)$$

σ_i denotes the usual spin taking $+1$ or -1. The energy parameters are determined by random numbers generated in the interval of $[-1,1]$. a and b are adjustable coefficients with positive values. The signs of energy parameters are taken to avoid frustrations according to the method of Kikuchi [12]. An arbitrary amount of frustrations can be incorporated into the no-frustration system by inverting an arbitrary number of signs of energy parameters, $\{J_{ij}\}$ and $\{K_i^{\alpha}\}$. Thus, we refer to a random system as a system with randomly selected energy parameters, a no-frustration system as that with parameters adjusted to be no frustration, and a partially frustrated system as that with some frustrations incorporated into a system with no frustration.

2.2. Numerical calculation

The lowest energy state was determined by the method described by Kikuchi [12]. The simulation was carried out by the Metropolis Monte Carlo algorithm [13]. A random spin state was generated at the beginning. As the definition of one step of the simulation, a spin is flipped followed by the Metropolis judgment based on the energy function, Eq. (1). This process was performed from the first spin to the Nth spin on the chain. In this paper, we also set $N = 20$ as in Kikuchi [12]. This routine was repeated a maximum of 1000 times. That is, the maximum number of the spin flips was 20 000. When the spin state of the system reached the lowest energy state, the simulation was stopped. The number of the steps required to reach the lowest energy state corresponds to the folding time of a protein. We ran 100 simulations with this procedure and analyzed the statistics of the results.

The simulations were carried out for a random system, a no-frustration system, and two partially frustrated systems. The numbers of inversions of signs in the energy parameters, i.e. the amounts of frustrations, were 20 and 30, respectively. We employed the same energy parameter set used by Kikuchi [12] for the no-frustration system.

3. Results

We first show the result of the simulation of the no-frustration system at $T = 4.5$ as an example. In this case, 97 of 100 trials reached the lowest energy sate. In other words, only three simulations could not attain the lowest energy state within 20 000 MC steps. The number of the Monte Carlo steps required to reach the lowest energy state defines the first passage time (FPT) to the state. We refer to the MC simulations that reached the lowest energy state as 'folded' or 'foldable' by analogy to proteins. The average FPT, i.e. the mean first passage time (MFPT), of folded simulations is 699.4, and the maximum FPT is 7812. The histogram of the FPT of this case is shown in Fig. 1. The majority of folded simulations show the FPT was less than 200 steps. The number of foldables may depend on the maximum number of MC steps in a simulation. For comparison, we performed the same simulations for some unfolded results with 10 000 times of the spin flips through the first to the Nth spins instead of 1000; unfolded simulations were still unfoldable. Thus, we could not find a simple relationship between the number of foldables and the number of the maximum MC steps, and the existence of unfoldables suggests the existence of unfoldable pathways or unfoldable regions on the energy landscape. The number of foldables of this system decreases at high and low temperature (T) as shown in Fig. 2 (the plot denoted by open rectangles). This result is interpreted to mean that in the higher temperature region the kinetics

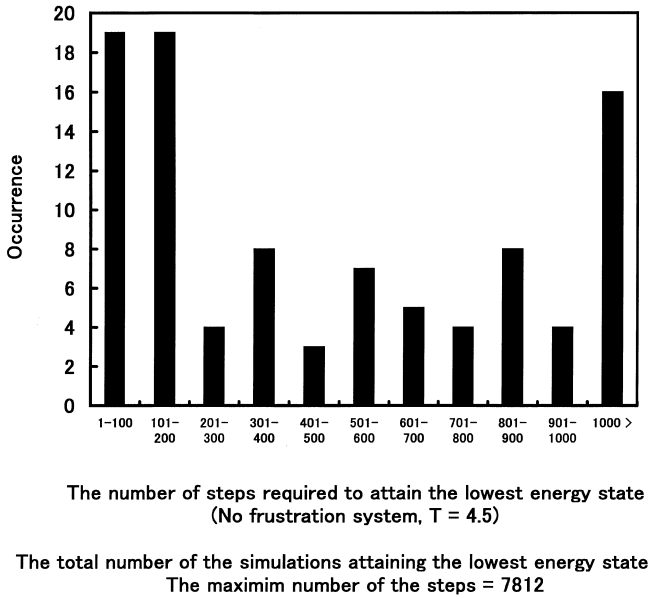

Fig. 1. The histogram of the number of steps required to reach the lowest energy state (FPT) for the no-frustration system at a simulation temperature 4.5. The total number of the simulations which attained the lowest energy state is 97 out of 100.

is described mainly by Levinthal entropy, whereas in the low temperature region several pathways are trapped in various non-lowest energy states.

In order to investigate the kinetics of the formation of the lowest energy state of the no-frustration system, we make a plot corresponding to an Arrhenius plot, i.e. the $1/T$ dependence of the logarithm of MFPT of foldables in Fig. 3 (denoted by open circles). At high temperature, i.e. the region where $1/T$ is less than 0.15 ($T \geqq 6.7$), the values of ln(MFPT) are staying at approximately 9.0. In the temperature region lower than $T = 6.7$ ($1/T \geqq 0.15$), the plot shows Arrhenius-like behavior as a plot of a protein folding kinetics until $1/T = 0.29$. In the region between $1/T = 0.29$ and 0.47, the plot becomes parabolic. This phenomenon also appears in the more realistic protein folding kinetics simulations [8,14] and folding experiments [15], and is considered to be attributed to the Gaussian distribution (random distribution) of the energy states on the energy landscape of a protein [Random Energy Model (REM)] [8], or temperature dependence of a hydrophobic interaction [14] between residues. This result in our simulation should be interpreted as the effect of the large dispersion of the energy states distribution as demonstrated in

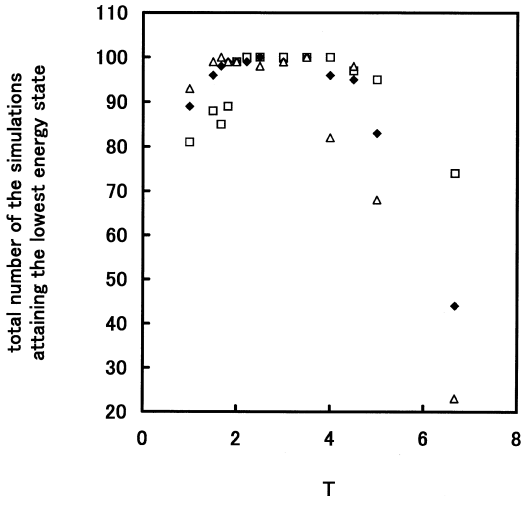

Fig. 2. Temperature (T) dependence of the total number of the simulations reaching the lowest energy state for the no-frustration system (□), for the frustrated system with $m = 20$ (◆) and for the frustrated system with $m = 30$ (△).

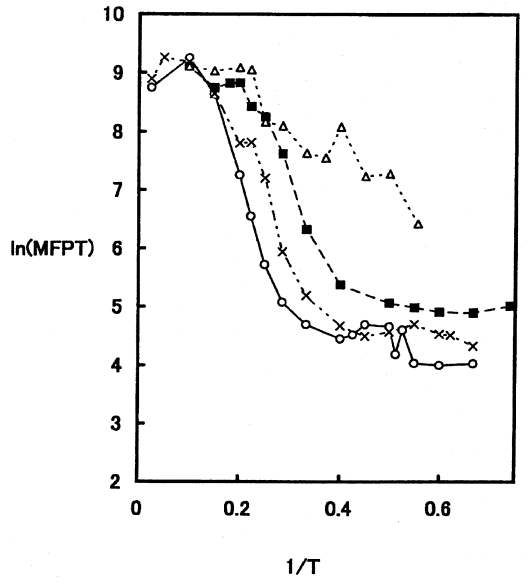

Fig. 3. $1/T$ dependence of the ln(MFPT) for the no-frustration system (○), for the frustrated system with $m = 20$ (×), for the frustrated system with $m = 30$ (■), and for the random system (△).

REM theory. Above $1/T = 0.47$ (below $T = 2.13$), ln(MFPT) rapidly decreases again and shows an almost constant value at approximately 4.0 accompanying the decrease of the number of foldables. (See also Fig. 2.) This result suggests that below $T = 2.13$ the system freezes into various minima on the potential energy surface and only initial states near the lowest energy state can 'freeze' into the lowest energy state with the ln(MFPT) value of approximately 4.0. In other words, the system cannot reach the self-averaged state in this very low temperature region. Yet the number of foldables is approximately 70–90 in this temperature range of Fig. 2. On the other hand, below $1/T = 0.47$, it is suggested from the Arrhenius and REM behavior that the system is self-averaged in the foldable regions on the energy landscape.

We next consider the effect of the frustrations. As mentioned in the method section, we incorporate the frustrations $m = 20$ and 30 in the no-frustrated sets of energy parameters, $\{J_{ij}\}$ and $\{K_i^{\alpha}\}$. The temperature dependence of the number of foldables and ln(MFPT) are represented in Figs. 2 and 3, respectively. The overall trends of those plots are almost the same as the no-frustration case. The interesting feature of the T dependence of the number of foldables in Fig. 2 is that the trend shifts from the high temperature region to thre low temperature region as the frustrations increase. That is, the lower frustrated system has a high number of the foldables at higher temperature. The frustrations decrease the number of foldables in the high temperature region, but increase it at low temperature regions. Fig. 3 also shows the $1/T$ dependence of ln(MFPT) of the frustrated systems [plotted by crosses ($m = 20$) and filled rectangles ($m = 30$), respectively]. At high temperature regions, the value of ln(MFPT) is approximately 9.0 in both systems. We already showed in the case of the no-frustration system that ln(MFPT) decreases at $T = 6.7$ and above ($1/T = 0.15$ and below) (Fig. 3) and almost all simulations starts to fold into the lowest energy state at this temperature (Fig. 2). For frustrated systems, the corresponding temperature is 4.5 ($1/T = 0.22$) in $m = 20$, and 3.5 ($1/T = 0.29$) in $m = 30$. That is, the temperature at which the decrease of ln(MFPT) occurs goes down as frustrations are added. Furthermore, the temperature at which the REM phase appears also decreases as the frustrations are incorporated, and the effect of REM itself becomes weak with the frustrations as shown in Fig. 3. The temperature at which freezing occurs is decreased by frustrations.

On the other hand, the T dependence of the number of foldables of the random system (Fig. 4) denotes that the plots are scattered especially at $T < 5.0$ ($1/T > 0.2$) suggesting the system is not self-averaged at this temperature region. The plots denoted by open triangles in Fig. 3, i.e. the $1/T$ dependence of ln(MFPT), also shows that the ln(MFPT) value of the random system at $1/T < 0.22$ ($T > 4.5$) are approximately 9.0, i.e. almost the same as the values of simulations at the high temperature. However, at $1/T > 0.22$ ($T < 4.5$), ln(MFPT) shows scattered plots with a decreasing trend. There is no Arrhenius region for the random system. These results denote that the random system at $T \gtrsim 4.5$–5.0 is self-averaged, but at $T < 4.5$, the system tends to be trapped in various states.

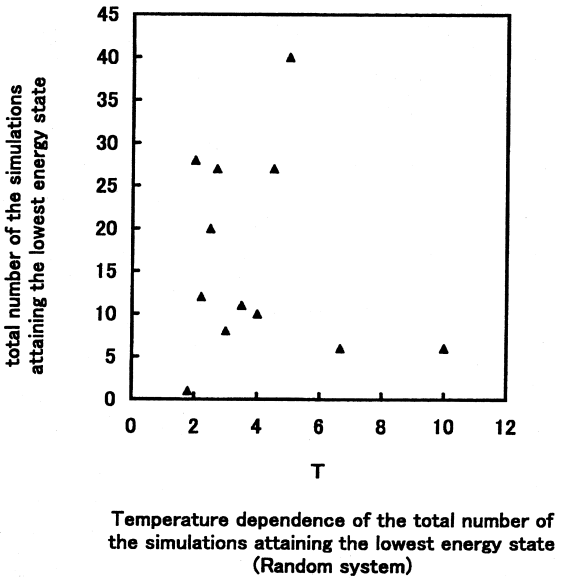

Fig. 4. Temperature (T) dependence of the total number of the simulations reaching the lowest energy state for the random system.

4. Discussion

From our results, it is concluded that the structure formation of the lowest energy state of the present spin system with no and some frustration(s) should be described by two factors: (i) the ratio of the number of foldables; and (ii) the rate of structure formation of the lowest energy state of foldables. Similar to other studies [16–18], in our systems an optimal temperature region for foldables exists, i.e. the number of foldables is maximized in this temperature region. This suggests that many simulations at higher temperature than this optimal temperature region were on unfoldable pathways with high energy on the energy landscape. In this case, the rate of foldables shows almost the same value for all simulations as shown in Fig. 3 disregarding the degree of frustrations. This phenomenon denotes the disappearance of the Arrhenius behavior indicating that the effect of a barrier is negligible. Thus, it is interpreted that the structure formation (folding) rate depends essentially on the Levinthal entropy of foldable pathways (probably entropy around the lowest energy state), i.e. the rate of the foldables is expressed as follows, where S_0 is the Levinthal entropy and k is the Boltzmann constant.

$$\tau_F \sim e^{S_0/k} \tag{2}$$

At very high temperature the thermally averaged shapes of the free energy profile of no frustration, frustrated and random systems are almost the same thus making the effect of barrier height disappear. The number of foldables tends to increase with decreasing temperature until reaching the optimal temperature region mentioned above. This corresponds to the temperature region in which ln(MFPT) of foldables exhibits the Arrhenius behavior as indicated in Fig. 3, implying that the folding rate of foldables is governed by a free energy barrier. Fig. 3 demonstrates that with decreasing temperature this effect appears first for the no-frustration case. That is, in the no-frustration system, the barrier is remarkable at relatively high temperature. As frustrations increase, the barrier effect appears distinctly in the lower temperature region. In other words, the rate in the no-frustration system becomes larger at relatively high temperature compared with partially-frustrated systems. At the temperature region where the barrier effect becomes noticeable, a very rapid decrease of τ_F is observed in these systems. Otherwise, the Levinthal entropy effect is large, and the value of τ_F remains large (especially as seen in the random system).

On the other hand, in a very low temperature region, many of the simulations fall into frozen states and cannot reach the lowest energy state as indicated in Fig. 2. Thus, the number of foldables decreases and the REM effect become dominant, i.e. the parabolic nature of ln(MFPT) becomes distinct. The interesting point is that the REM effect in the no-frustration system appears at higher temperature than in partially frustrated systems (Fig. 3). Furthermore, the effect becomes weak (the curvature of the parabolic part becomes small) as frustrations are added. The effect of the ruggedness parameter $(\Delta E)^2$ in REM [8] is remarkable especially for the no-frustration system, where the dispersion of energy levels of the systems is large [8]. Our results demonstrate that

the energy levels are dispersed around the lowest energy state in the no-frustration system. Increasing the amount of frustrations causes the REM effect to become small. This effect also shows up in the increased number of foldables in the lower temperature region for the partially frustrated systems as shown in Fig. 2, in which the energy landscape becomes less rugged compared with the case of no-frustration system. Therefore, in the very low-temperature region, ln(MFPT) of the no-frustration system is larger than that of a frustrated system because of the parabolic nature of the REM effect [compare the values of ln(MFPT) of the no-frustration and the partially frustrated system with $m = 20$ at the temperature region $0.4 < 1/T < 0.5$ in Fig. 3].

Hence, the behavior of the rate of lowest energy state formation can be interpreted in the model of the free energy property as a function of the order parameter (Fig. 5). At higher temperature, the free energy profile is flat over order parameter values. At decreasing temperatures the free energy curve possesses a barrier between the native and disordered states. In the very low temperature region, ruggedness of the free energy profile becomes remarkable, and then the REM property of the folding rate appears.

Thus, summarizing the present discussion, the folding rate of the present systems should be formally expressed as follows:

$$t \sim \tau_F(T)/P(T) \tag{3}$$

Here t is an overall folding rate and $\tau_F(T)$ denotes the folding rate of foldables, i.e., MFPT. $P(T)$ denotes the ratio of foldables to total number of molecules. This result is derived from the simulations of the spin system defined by Eq. (1). However, it is possible that a similar situation occurs in actual protein folding kinetics. In an actual protein, the folding rate with the form of Eq. (2), i.e. the rate controlled by the Levinthal entropy, becomes too small to be observed for the reason stated in the Levinthal paradox [1], and only the rate governed by the free energy barrier is observed. The parabolic nature of folding rate at the very low temperature was observed in several kinetic experiments in relation to cold denaturation [19,20], but the origin of this phenomenon is controversial. The folding simulations of lattice protein models sometimes show this effect which is interpreted by the REM model [8,16,21] as discussed in the present paper. On the other hand, Chan and Dill [14] attributed this effect to the specific temperature dependence of interaction parameters because of hydrophobic interactions [22,23]. As far as the present model is concerned, the specific temperature dependence in the interaction parameters is not taken into account. It is natural to think that the cold denaturation phenomenon involves the solvation effect including hydrophobic interactions, but it might reflect the ruggedness of the potential energy surface of a protein with water molecules. We think that the temperature dependence of inter-

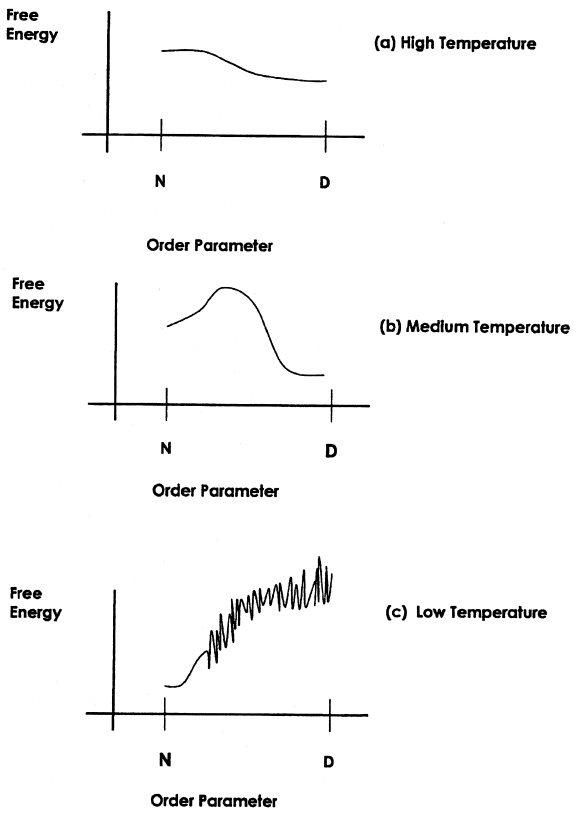

Fig. 5. Hypothetical free energy profile of the present spin systems as functions of the order parameter at (a) high temperature, (b) medium temperature, and (c) low temperature.

action parameters will be interpreted by the theory similar to REM.

The present spin model is very simple, and therefore it is possible to analyze in detail the effect of frustrations on the structure formation of the lowest energy spin state which is a model of the protein folding. Thus, even the analysis of the Levinthal entropy control state is also possible. We found out within the present model that at a certain temperature region the folding rate is mainly governed by the free energy barrier especially in the no-frustration system (because the ratio of foldables is almost 1 at this temperature region) and the folding time becomes very short. Furthermore, the transition of Levinthal entropy control to the barrier control occurs at relatively high temperature in the no-frustration system. This is a crucial difference of a rapidly foldable system from a largely frustrated or random system. We believe that the same situation appears in actual protein folding.

We have a further question: how can a protein system produce a free energy barrier at a certain temperature region to exhibit the two-state transition, in other words, how can a denatured state of a protein (or a spin model) can constitute an equilibrium ensemble of many of states which are convertible to the native state at a certain temperature region? We are currently studying this problem using the present spin systems in order to obtain further insight into the protein folding mechanism.

References

[1] C. Levinthal, J. Chim. Phys. 85 (1968) 44.
[2] J.N. Onuchic, P.G. Wolynes, Z. Luthey-Schulten, N.D. Socci, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 3626.
[3] P.E. Leopold, M. Montal, J.N. Onuchic, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 8721.
[4] J.D. Bryngelson, P.G. Wolynes, J. Phys. Chem. 93 (1989) 6902.
[5] N. Go, Annu. Rev. Biophys. Bioeng. 12 (1983) 183.
[6] J. Wang, J.G. Saven, P.G. Wolynes, J. Chem. Phys. 105 (1996) 11276.
[7] S.S. Plotkin, J. Wang, P.G. Wolynes, J. Chem. Phys. 106 (1997) 2932.
[8] J. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, Proteins 21 (1995) 167.
[9] D.K. Klimov, D. Thirumalai, Proteins 26 (1996) 411.
[10] D.K. Klimov, D. Thirumalai, J. Chem. Phys. 109 (1998) 4119.
[11] A. Sali, E. Shakhnovich, M. Karplus, J. Mol. Biol. 235 (1994) 1614.
[12] T. Kikuchi, Biophys. Chem. 65 (1997) 109.
[13] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, J. Chem. Phys. 21 (1953) 1087.
[14] H.S. Chan, K.A. Dill, Proteins 30 (1998) 2.
[15] T.E. Creighton, in: R.H. Pain (Ed.), Mechanisms of Protein Folding, Oxford University Press, Oxford, 1994, p. 1.
[16] V.I. Abkevich, A.M. Gutin, E.I. Shakhnovich, J. Chem. Phys. 101 (1994) 6052.
[17] R. Miller, C.A. Danko, M.J. Fasolka, A.C. Balazs, H.S. Chan, K.A. Dill, J. Chem. Phys. 96 (1992) 768.
[18] H.S. Chan, K.A. Dill, J. Chem.Phys. 100 (1994) 9238.
[19] B. −L. Chen, W.A. Baase, J.A. Schellman, Biochemistry 28 (1989) 691.
[20] S.E. Jackson, A.R. Fersht, Biochemistry 30 (1991) 10428.
[21] N.D. Socci, J.N. Onuchic, J. Chem. Phys. 101 (1994) 1519.
[22] D.O.V. Alonso, K.A. Dill, Biochemistry 30 (1992) 5974.
[23] K.A. Dill, D.O.V. Alonso, K. Hutchinson, Biochemistry 28 (1989) 5439.