Protein Folding under Mediation of Ordering Water: an Off-Lattice Gō-Like Model Study *

ZUO Guang-Hong(左光宏)[1,2], HU Jun[1,3](胡钧), FANG Hai-Ping(方海平)[1,2]**

[1] *Shanghai Institute of Applied Physics, Chinese Academy of Sciences, PO Box 800-204, Shanghai 201800*

[2] *T-Life Research Center, Fudan University, Shanghai 200433*

[3] *Bio-X Life Science Research Center, College of Life Science and Biotechnology, Shanghai Jiao Tong University, Shanghai 200030*

(Received 13 June 2007)

Water plays an important role in the structure and function of biomolecules. Water confined at the nanoscale usually exhibits phenomena not seen in bulk water, including the ice-like ordering structure on the surfaces of many substrates. We investigate the behaviour of protein folding in which the proteins are assumed in an environment with ordering water by using of an off-lattice Gō-like model. It is found that in the physiological temperature, both the folding rate and the thermodynamic stability of the protein are greatly promoted by the existence of ordering of water.

PACS: 87.14.Ee, 87.15.Cc, 87.64.Aa

How proteins fold to their functional conformations is one of the most important and unsolved problems in molecular biology. There have been extensive studies in exploring the answer to this important question both experimentally and theoretically.[1−6] Most of those studies have been focused on protein folding *in vitro*. It has been well recognized that the protein folding behaviour *in vivo* is much complex than that *in vitro* because the proteins are in an environment full of water molecules, lipids, carbohydrates and other biological molecules.[7] Consequently, the folding of proteins *in vivo* usually take place under heterogeneous potentials in confined spaces,[8] whose sizes are at the nanometre scale.[9−11] The folding behaviour of protein *in vivo* should be different from that *in vitro*.

In recent years, it becomes recognized that water plays an important role in the structure and function of proteins.[12−15] The biological water described by Despa[16] is the water molecule confined inside or on the surface of biomolecules. Recently, it has been recognized that water confined on the micro- and nanoscale usually shows unusual properties never seen in its bulk counterpart. By using a scanning polarization microscopy, Hu and his co-workers observed that water condensed into solid-like polygonal domains in angular epitaxial relationship with the underlying muscovite mica surface at room temperature.[17,18] Later, this ice-like ordering structure has been observed on many solid surfaces experimentally.[19,20] First principle molecular dynamics simulation was used to analyse this ordering structure. It was found that water condense into a fully connected two-dimensional hydrogen-bond network, like the structure of ice, at monolayer coverage.[21] Very recently, molecular dynamics simulations showed that such an ice-like hydrogen-bond network even reached the third layer from the surface.[22,23]

It is possible that some of the water molecules confined in cellular environment exhibit ordered structure. In fact, it has been shown experimentally that some water molecules on the surface of DNA molecules are ordered, by using the femtosecond-resolved fluorescence up-conversion technique.[13,14,24] This raises the intriguing question of how the ordering water, when it exists in cell, contributes to the folding of proteins.

In this Letter, we present our study towards this direction by numerical simulating the folding of protein Chymotrypsin Inhibitor Ⅱ (CI2, PDB Code: 1coa) in an environment with ordering water (EWOW). CI2 is a well-known two-state folder and usually regarded as a model protein in the simulational studies of protein folding.[25−27] As a comparison, the folding of the model protein in the environment without the ordering water (EOOW) is also simulated. It is found that, in the physiological temperature, the model protein folds faster in EWOW than in EOOW. The folding transition temperature of the protein in EWOW is higher than that of the protein in EOOW. In the free energy for the protein folding in EWOW, the barrier from the denature state to the native state disappears when the temperature is around the body temperature of human being. Those results indicate that both the folding rate and the thermodynamic stability of the protein are promoted by the existence of ordering water.

* Supported by the National Natural Science Foundation of China under Grant Nos 10474109 and 10674146, and the Chinese Academy of Sciences.

** To whom correspondence should be addressed. Email: fanghaiping@sinap.ac.cn

©2007 Chinese Physical Society and IOP Publishing Ltd

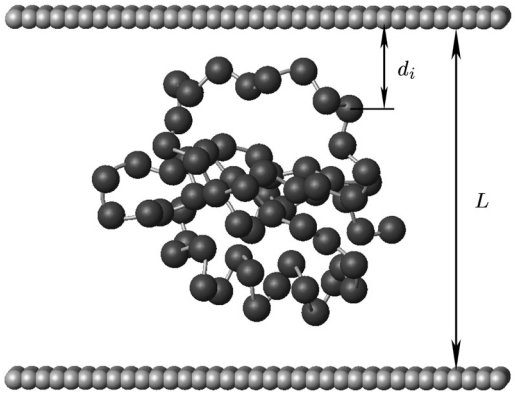

Fig. 1. Schematic diagram for protein folding confined in the space between two parallel slabs. In this model, the connected black beads, which are confined by two grey-bead slabs, represent the residues of the protein, the water molecules are implicit and not shown here. L denotes the distance of the two parallel slabs and d_i denotes the distance of the ith residue to its nearer slab.

The model used in this work is shown in Fig. 1. The protein is modelled as connected beads with Gō-like potential,[28] which has been widely used to describe the interactions between residues in the modelling of proteins.[9−11,25−27,29] The folding of the protein is confined in the space between two $L = 68.4\,\text{Å}$ apart parallel hard slabs and the ordering water formed on the inner flanks of this two slabs. A repulsion interaction between the residue and the slab is assumed when their distance is less than $\sigma_0 = 2\,\text{Å}$, the radius of hard core of the residue, which is realized by a Lennard–Jones form:[10,11]

$$V_r = \sum_i 50\Big[\Big(\frac{\sigma_0}{d_i}\Big)^4 - 2\Big(\frac{\sigma_0}{d_i}\Big)^2 + 1\Big]\Theta(\sigma_0 - d_i), \quad (1)$$

where d_i is the distance between the ith residue and its nearer slab, and $\Theta(x) = 1$ when $x > 0$, 0 otherwise. The random coil conformation is used as the initial conformation and the leap-frog algorithm[30] is used to integrate the equations of the Langevin dynamics. Considering that the ordering water has a weaker moving fluctuation than that of the bulk water, a factor $\lambda_i = 1 - e^{-d_i/d_0}$ is timed on the random force item of the Langevin equation ($\Gamma(t)$ of Eq. (2) of Ref. [25]), to realize the effect of the ordering water.[16,17,21−23] Here we have taken into account the fact that water is not ordered in bulk so that we assume that the degree of the ordering of the water near the slab surface decreases quickly with respect to the distance from the slab. We assume that $d_0 = 4.34\,\text{Å}$, so that $\lambda_i = 0.99$ for $d = 10\,\text{Å}$. The other choice of the value d_0 slightly different from 4.34 Å does not change the conclusion which obtained in this study. The thermodynamic quantities are calculated by the weighted histogram analysis method (WHAM).[31,32] The detailed descriptions of the other settings and parameters, including the Gō-like potential, Langevin dynamics and WHAM, can be referred to Ref. [25].

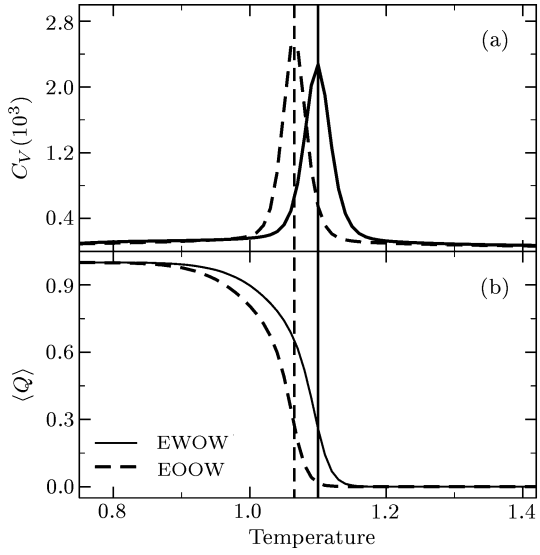

Fig. 2. Thermodynamical properties for protein folding in the environment with (sold lines) and without (dash lines) the ordering water (EWOW and EOOW). (a) The specific heat as a function of temperature. (b) The temperature dependence of average formed fraction of native contacts, Q. Two vertical lines denote the folding transition temperature of protein in two environments, respectively.

Base on the Gō-like model, the folding of the model protein CI2 confined in the space between two parallel slabs, with or without the ordering water on the inner flanks of the slabs, are simulated. The thermodynamic properties of the protein folding in these two different environments are described by the specific heat and the average formed fraction of native contacts, Q, versus the temperature, which are shown in Fig. 2. The sharp unimodal specific heat curves (see Fig. 2(a)) and the sigmoidal transition curves (see Fig. 2(b)) indicate that the protein folds in a two-state manner and the folding transition is very cooperative in each environment. The behaviour is similar to the experimental result *in vitro*[33] and to the simulation by the classical Gō-like model,[25−27] which simulates the protein folding in free space and is a model of the folding *in vitro*. However, the folding transition temperatures T_f in different environments is different for the protein folding. Here T_f is defined at the peak of specific heat (see Fig. 2(a)). Numerically we find $T_f = 1.10$ for the protein folds in EWOW, $T_f = 1.07$ for the protein folds in EOOW. They are both higher than those of the protein folding simulated with classical Gō-like model, where $T_f = 1.06$ as the same as Ref. [25]. Since a higher folding transition temperature indicates higher thermodynamic stability, both the mediation of the ordering water and the confinement promote the thermodynamic stability of the protein. We note that the contribution of the confinement to the thermodynamic stability had been reported in

the literature.[9−11] Here, we show that the folding transition temperature further significantly increases by the mediation of the ordering water.

The folding kinetics of the protein is usually described by the mean first passage time (MFPT), which denotes the average time required to fold a protein starting from a random coil conformation ($Q \leq 0.3$) to the native state ($Q \geq 0.9$), at seven temperatures in this work. The results are shown in Fig. 3. To obtain a reasonable folding time of the protein, 100 runs have been processed for each point in this figure. As shown in Fig. 3, there is a crosspoint for these two MFPT curves at $T_c \sim 0.85$. When the temperature $T < T_c$, the protein folds faster in EOOW than that in EWOW, and *vice versa*. Multiplied the ratio of T_f of the experiment *in vitro* to that of the classical Gō-like model,[25,33] the temperature value in our simulation scale could be converted into the value in the Kelvin scale (for the protein CI2(1coa), $T_f = 338\,\mathrm{K}$ in the experiment *in vitro*,[33] and $T_f = 1.06$ in simulation of classical Gō-like model,[25] which is the simulation of the protein folding *in vitro*, then $T_f^{\mathrm{exp}} \approx 317.3 \times T_f^{\mathrm{sim}}$). With this ratio, the temperature of the crosspoint and the maximum temperature shown in Fig. 3 are about 265 K and 330 K, respectively. This temperature range covers the range of physiological temperature for most of living beings. Thus, the mediation of the ordering water speeds the protein folding in the physiological temperature.

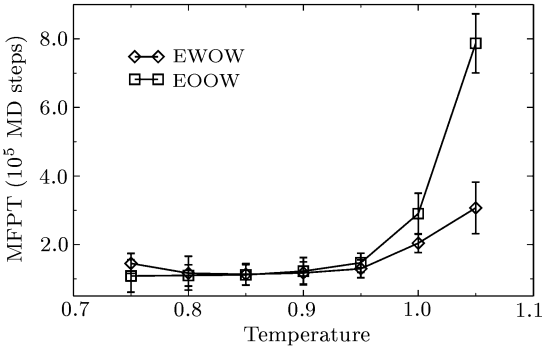

Fig. 3. Mean first passage time (MFPT) of the protein folding in the environment with and without the ordering water (EWOW and EOOW) at seven different temperatures.

Recently, the funnel-like energy theory has been widely accepted as the framework to describe the folding of proteins.[5,34] Here the free energy landscape of protein folding is calculated by the WHAM[25−27,29] and shown in Fig. 4. There are two free energy landscapes for the protein folds in EWOW and EOOW at $T = 0.99 \approx 37^\circ\mathrm{C}$ in Fig. 4. With the funnel-like landscape theory, the folding properties of a protein can be described by the relative height between states in the free energy landscape.[5] In order to compare the folding properties of the protein in EWOW and EOOW, the minima of the native states for the two curves are assumed equal. As shown in Fig. 4, there are two minima in the free energy landscape of the protein folding in EOOW, which correspond to the denature state and native state, respectively. A barrier, which corresponds to the transition state, exists between these two minima. Such a free energy landscape indicates that the protein folds in a two-state manner in EOOW at this temperature. Therefore, the folding process of this protein in this environment can be described as the hopping from the denatured state to the native state without any accumulation of intermediate states, and the barrier becomes the main obstacle of the folding at this temperature. However, with the mediation of the ordering water, both the corresponding minimum of the denature state and the corresponding barrier of the transition state almost disappear. This makes the hopping from the denature state to the native state easier. That is, with the mediation of the ordering water, the folding of the protein becomes faster.

Another important observation is the difference of the free energies in the interval between $Q_t < Q < Q_n$, where $Q_t \approx 0.4$ corresponds to the transition state of the free energy landscape for the protein folds in EWOW, and at $Q_n \approx 0.9$, corresponding to the native state, the free energy landscapes have the minima for both the cases. The free energies for both the cases decrease monotonically, while the free energy for protein folding in EWOW decreases faster than that in EOOW. Therefore, the unfolding of the protein is more difficult from the native state to the denature state in EWOW. The stability of the protein is improved by the ordering water at this temperature.

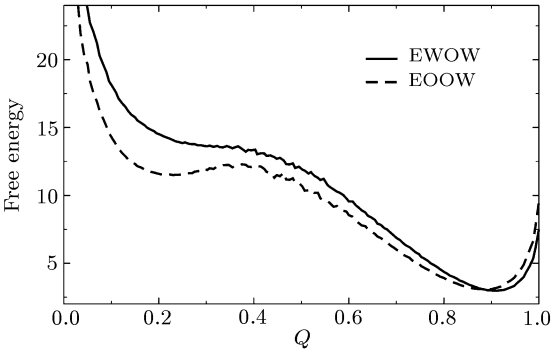

Fig. 4. Free energy landscape of the protein in the environment with (solid line) and without (dash line) the ordering water (EWOW and EOOW) at $T = 0.99$, which is about the 37°C in Celsius scale.

In conclusion, we have studied the folding of protein CI2 under the mediation of the ordering water by using of the off-lattice Gō-like model, and compared it with the folding of the protein under the same conditions but without the ordering water. Our results

show that the ordering water plays an important role in the protein folding. It speeds the protein folding at the physiological temperature and makes the protein more stable at high temperature. It is noted that though the confinement is helpful to the thermodynamic stability and folding rate of the protein, the existence of the ordering water on the inner flank of the confinement space further significantly improves the thermodynamic stability and folding rate of the protein. Therefore, if there is ordering water formed on the surface of the cellular components in tissues and cells, the proteins will be more stable and fold faster *in vivo* than those *in vitro*.

We thank Bo Zhou and Yizhen Lu for helpful discussion.

References

[1] Anfinsen C B 1973 *Science* **181** 223
[2] Fersht A R 1995 *Curr. Opin. Struct. Biol.* **5** 79
[3] Brooks III C L 1998 *Curr. Opin. Struct. Biol.* **8** 222
[4] Shakhnovich E 2006 *Chem. Rev.* **106** 1559
[5] Onuchic J N, Luthey-Schulten Z and Wolynes P G 1997 *Annu. Rev. Phys. Chem.* **48** 545
[6] Li M S, Klimov K K and Thirumalai D 2004 *Phys. Rev. Lett.* **93** 268107
[7] Ellis R J 2001 *Curr. Opin. Struct. Biol.* **11** 114
[8] Minton A P 1992 *Biophys. J.* **63** 1090
[9] Klimov D K, Newfield D and Thirumalai D 2002 *Proc. Natl. Acad. Sci. USA* **99** 8019
[10] Takagi F, Koga N and Takagi S 2003 *Proc. Natl. Acad. Sci. USA* **100** 11367
[11] Xu W X, Wang J and Wang W 2005 *Proteins* **61** 777
[12] Papoian P A, Ulander J, Eastwood M P, Luthey-Schulten Z and Wolynes P G 2004 *Proc. Natl. Acad. Sci. USA* **101** 3352
[13] Pal S K, Peon J and Zewail A H 2002 *Proc. Natl. Acad. Sci. USA* **99** 1763
[14] Pal S K, Zhao L and Zewail A H 2003 *Proc. Natl. Acad. Sci. USA* **100** 8113
[15] Chaplin M 2006 *Nat. Rev. Mol. Cell Biol.* **7** 861
[16] Despa F 2005 *Ann. N.Y. Acad. Sci.* **1066** 1
[17] Hu J, Xiao X D, Ogletree D F and Salmeron M 1995 *Science* **268** 267
[18] Hu J, Xiao X D, Ogletree D F and Salmeron M 1995 *Surf. Sci.* **344** 221
[19] Reedijk M F, Arsic J, Hollander F F, de Vries S A and Vlieg E 2003 *Phys. Rev. Lett.* **90** 066103
[20] Aarts I M, Pipino A C, Hoefnagels J P, Kessels W M and van de Sanden M C 2005 *Phys Rev Lett* **95** 166104
[21] Odelius M, Bernasconi M and Parrinello M 1997 *Phys. Rev. Lett.* **78** 2855
[22] Zhou B and Fang H P (in preparation)
[23] Leng Y S and Cummings P T 2006 *J. Chem. Phys.* **124** 074711
[24] Pal S K and Zewail A H 2004 *Chem. Rev.* **104** 2099
[25] Zuo G H, Wang J and Wang W 2006 *Proteins* **63** 165
[26] Kaya H and Chan H S 2003 *J. Mol. Biol.* **326** 911
[27] Clementi C, Nymeyer H and Onuchic J N 2001 *J. Mol. Biol.* **298** 937
[28] Ueda Y, Taketomi H and Gō N 1975 *Int. J. Pept. Res.* **7** 445
[29] Zuo G H, Zhang J, Wang J and Wang W 2005 *Chin. Phys. Lett.* **22** 1809
[30] Allen M P and Tildesley D J 1987 *Computer Simulation of Liquid.* (Oxford: Clarendon)
[31] Ferrenberg A M and Swendsen R H 1989 *Phys. Rev. Lett.* **63** 1195
[32] Shea J E and Brooks III C L 2001 *Annu. Rev. Phys. Chem.* **52** 499
[33] Jackson J E et al 1993 *Biochemistry* **32** 11259
[34] Dill K A and Chan H H 1997 *Nature Strut. Biol.* **4** 10