ELSEVIER

Physica D 107 (1997) 322–325

PHYSICA D

Statistical mechanics of correlated energy landscape models for random heteropolymers and proteins

Steven S. Plotkin *, Jin Wang, Peter G. Wolynes [1]

School of Chemical Sciences, University of Illinois, Urbana, IL 61801, USA

Abstract

We study the role of correlations in the energy landscape of heteropolymers and proteins, specifically their role in the glass transition in random heteropolymers and the folding transition in minimally frustrated proteins. In the context of the glass transition, a correlated landscape results in a more gradual freezing into basins of extensive entropy, while not completely destroying the first-order jump in the order parameter until below a certain density. Quantities such as the glass transition temperature and the probability distribution of overlaps q are quantitatively similar to the results for an uncorrelated landscape or random energy model (REM), while the number of searchable basins at the glass transition (the Levinthal search) is significantly modified.

For proteins, correlations provide a way to induce a funnel topography onto the energy landscape by the selection of a sequence with a particularly low energy configuration. The folding transition is weakly first-order. The position of the transition state ensemble in the model is in accord with recent experimental results on denaturant effects on kinetics of small proteins.

Keywords: Proteins; Random heteropolymers; Spin-glasses; Folding transition; Disordered systems

1. The glass transition in heteropolymers

We study the effects of correlations in the energy landscape on the glass transition in a random heteropolymer [1]. If two states are configurationally similar as defined by an order parameter q measuring the fraction of identical contacts in each configuration, then their energies tend to be similar. A simple contact hamiltonian $\mathcal{H} = \sum \varepsilon_{ij}\sigma_{ij}$ with gaussian random contact energies ε_{ij} and $\sigma_{ij} = 0$ or 1, yields a joint probability distribution for the energies of two states which is identical to that obtained in a generalized random energy model (GREM) of Derrida and Gardner [2], where all states are ultrametrically organized. This means that the bond averaged free energy $-T\langle \log Z(T)\rangle$ obtained through the GREM formalism [2] will be accurate to second-order in the partition function of the system (i.e. $\langle Z(T)\rangle$ and the correlations $\langle Z(T)Z(T')\rangle$ are the same for both the GREM model and the contact hamiltonian).

In order to obtain the free energy, we have explicitly calculated the configurational entropy of a polymer subject to either few or very many topological constraints, and interpolated between the two regimes. At low q the entropy is reduced from that of the unconstrained system by the formation of cross-links, essentially following the Flory formula for vulcanization [3], but with the following modifications: (1) finite

* Corresponding author. Tel.: 217 333 8710; fax: 217 244 5984; e-mail: plotkin@odin.scs.uiuc.edu.

[1] E-mail: wolynes@aries.scs.uiuc.edu.

0167-2789/97/$17.00 Copyright © 1997 Elsevier Science B.V. All rights reserved
PII S0167-2789(97)00100-0

size confinement effects are taken into account; (2) combinatorial entropy is accounted for in choosing the set of constraining bonds from the total possible bonds; (3) we also consider the fact that a number of bonds $\propto (1-q)$ cannot have been formed, which further reduces the entropy.

When the number of bonds exceeds the number of monomers in the system, the entropy is then calculated by considering the combinatorics of discrete sections of the polymer chain "melting out" from the frozen ($q = 1$) three-dimensional structure of a reference state. Knowing the entropics and energetics allows the free energy to be determined within the GREM formalism.

For collapsed polymers the system exhibits a random first-order glass transition below a temperature T_g, which confines non-glassy polymer dynamics to a basin of states of extensive entropy $\approx 1/4$ of the total entropy. This phase transition is representative of a wide universality class of spin-glass transitions in systems that lack special symmetries [5]. The number of basins to be searched below T_g is reduced from the uncorrelated landscape value ($\exp S_0$ in the REM to $\approx \exp(0.7S_0)$ in the GREM for collapsed 27-mers, where S_0 is the total unconstrained entropy), which decreases the Levinthal search time [4] for proteins with glassy dynamics. The number of basins at T_g is a sensitive function of density. For packing fractions of $\cong 0.85$, there are only $\approx \exp(0.25S_0)$ basins to search through. Similar reductions in the number of basins searched to escape from a meta-stable state at the glass temperature have been obtained in replica variational [6] and correlated landscape kinetic treatments [7].

In the GREM, the glass temperature at which freezing begins is quantitatively similar to the uncorrelated landscape temperature T_{rem}, and is given by

$$T_g \cong \left(\frac{Q_o \varepsilon^2}{2s_{LEV}}\right)^{1/2}, \tag{1}$$

where s_{LEV} is the Levinthal entropy per monomer, ε measures the roughness of the landscape, and Q_o is the minimum similarity in a confined basin (all states in the basin have overlaps $\geq Q_o$). For temperatures less than T_g the melted basin gradually shrinks to a single frozen state at a temperature considerably lower than T_{rem}.

In investigating the density dependence of the glass temperaure, it was found that below a tri-critical density (of packing fraction $\eta_c \cong 0.85$ for a 27-mer) the order parameter $q(x)$ becomes second-order continuous. However, characterizations such as the bond averaged probability of overlap $P(q)$ at temperature T remain quantitatively similar.

2. The folding transition in proteins

To model the effect of minimal frustration [8] in proteins or designed heteropolymers, we simply require the energy of a particular state to be lower than the average by a "stability gap" energy $\delta\varepsilon_n$ – an additional energy parameter in the theory. Energetic correlations between states induce a funnel [9] topography around a given state of low energy [10]. For sufficient stability gap energies compared to the roughness ($\delta\varepsilon_n/\varepsilon \gtrsim \sqrt{2}$), the folding temperature of the protein is above its glass transition temperature, and the model may be analyzed in the replica-symmetric regime where glassy trapping does not play a role.

A simple model coupling polymer density to bond formation was introduced to obtain a free energy surface (as a function of the number of native contacts and the total number of contacts) which captures the qualitative features of simulations [11] (see Fig. 1). The corresponding states principle [12] is then used to compare the properties of transition states of real proteins with those of (smaller length) lattice models. In the model, the folding transition is weakly first-order with a small barrier. For small proteins (e.g. λ-repressor), the predicted free energy barrier for folding is a few $k_B T$, which is small compared to the entropic barrier of $\approx 21 k_B T$. The barrier results from the subtle incomplete cancellation of entropic loss and negative energy gain. The transition state ensemble has about $1/4$ of the unconstrained molten globule entropy and contains about $1/2$ the native contacts. While the position of the barrier in coordinate Q agrees with that of small proteins [13], the experimental barrier positions for larger proteins [14,15] appear smaller than those

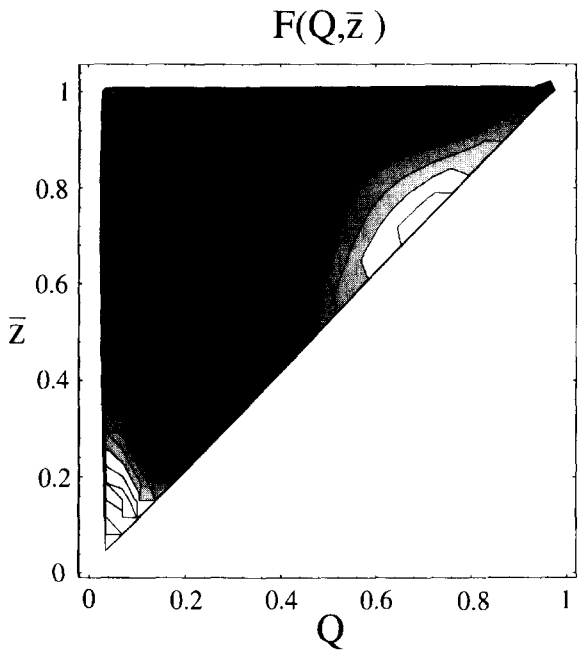

Fig. 1. Free energy surface at the folding temperature for the 27-mer, with typical roughness, gap and hydrophobicity (dark areas are deeper in free energy). The surface has a double-well structure with a transition state bottleneck at $Q^* \cong 1/2$ and barrier height $\Delta F \cong 3k_BT$.

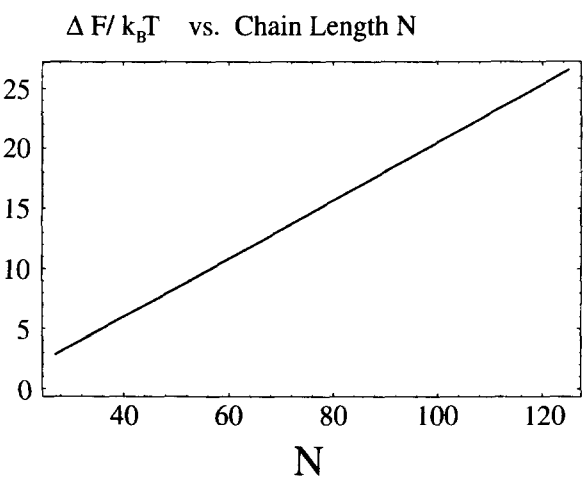

Fig. 2. Free energy barrier height ΔF in units of k_BT, as a function of sequence length N.

predicted by the theory, indicating the possibility of local regions of the chain rearranging separately via a *foldon* mechanism [16], as opposed to homogeneous mean-field behavior.

The mean-field theory [10] predicts a folding barrier rising nearly linearly with N (see Fig. 2), while other theories considering either nucleation growth of a contiguous native core [18], or free energy fluctuations analagous to those in Potts glasses [19], give barriers scaling as $N^{2/3}$ or $N^{1/2}$, respectively. Experimental results will eventually reveal which of these mechanisms is dominant for proteins of various sizes, hydrophobicity, intrinsic roughness, and stability gap [20]. Other theories that allow local regions of the chain to rearrange separately may also be significant [16,17].

Acknowledgements

This work was supported by NIH grant 1RO1 GM44557, and NSF grant DMR-89-20538.

References

[1] S.S. Plotkin, J. Wang and P.G. Wolynes, Phys. Rev. E 53 (1996) 6271.
[2] B. Derrida and E. Gardner, J. Phys. C 19 (1986) 2253–2274.
[3] P.J. Flory, J. Am. Chem. Soc. 78 (1956) 5222.
[4] C. Levinthal, in: Mossbauer Spectroscopy in Biological Systems; Proc. a Meeting Held at Allerton House, Monticello, IL (University of Illinois Press, Urbana, IL, 1969) 22.
[5] D.J. Gross, I. Kantor and H. Sompolinsky, Phys. Rev. Lett. 55 (1985) 304; T.R. Kirkpatrick and P.G. Wolynes, Phys. Rev. B 36 (1987) 8552–8564.
[6] S. Takada and P.G. Wolynes, Phys. Rev. E 55 (1997) 4562–4577.
[7] J. Wang, S.S. Plotkin and P.G. Wolynes, J. de Phys. I 7 (1997) 395–421.
[8] N. Gō , Ann. Rev. Biophys. Bioeng. 12 (1983) 183–210.
[9] J. Bryngelson, J.O. Onuchic, N.D. Socci and P.G. Wolynes, Proteins: Structure, Function, and Genetics 21 (1995) 167–195.
[10] S.S. Plotkin, J. Wang and P.G. Wolynes, J. Chem. Phys., 106 (1997) 2932–2948.
[11] A. Kolinski, A. Godzik and J. Skolnick, J. Chem. Phys. 98 (9) (1993) 7420; N.D. Socci and J.N. Onuchic, J. Chem. Phys. 101 (1994) 1519–1528; K.A. Dill, S. Bromberg, K. Yue, K.M. Fiebig, D.P. Yee, P.D. Thomas and H.S. Chan, Protein Science 4 (1995) 561–602; A. Sali, E. Shakhnovich and M. Karplus, J. Mol. Biol. 235 (1994) 1614.
[12] J.N. Onuchic, P.G. Wolynes, Z. Luthey-Schulten and N.D. Socci, Proc. Natl. Acad. Sci. USA 92 (1995) 3626–3630.

[13] G.S. Huang and T.G. Oas, Proc. Natl. Acad. Sci. USA 92 (1995) 6878.
[14] S.E. Jackson and A.R. Fersht, Biochem. 30 (1991) 10428.
[15] T. Pascher, J.P. Chesick, J.R. Winkler and H.B. Gray, Science 271 (1996) 1558.
[16] A. Panchenko, Z.A. Luthey-Schulten and P.G. Wolynes, Proc. Natl. Acad. Sci. USA 93 (1996) 2008.
[17] S. Takada and P.G. Wolynes, Prog. Theor. Phys. Suppl., in press.
[18] A.E. Finkelstein and A.Ya. Badretdinov, Folding and Design 2 (1997) 115–121.
[19] D. Thirumalai, J. de Phys. I 5 (1995) 1457–1467.
[20] P.G. Wolynes, Proc. Natl. Acad. Sci. USA 94 (1997) 6170–6175.