Dynamics of Random Hydrophobic-Hydrophilic Copolymers with Implications for Protein Folding

D. Thirumalai,[1] V. Ashwin,[2] and J. K. Bhattacharjee[3]
[1]*Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742*
[2]*Department of Physics, Indian Institute of Technology, Kanpur 208016, India*
(Received 11 March 1996)

We study the dynamics of a single amphiphilic random copolymer chain. The dynamical treatment shows that in addition to the collapse temperature, T_Θ, there exists a scale dependent glass transition temperature $T_G(l)$. Distinct scenarios for the glassy behavior emerges depending on the relative values of T_Θ and $T_G(l)$. The possible implications of our results for protein folding, including an estimate of the dependence of the ratio of the folding transition temperature to $T_G(l)$ on the length of the chain, are sketched. [S0031-9007(96)01832-7]

PACS numbers: 61.41.+e, 64.70.Pf

Inspired at least partly by the protein folding problem, a class of disordered heteropolymer models have been introduced over the past several years [1,2]. The objective of these studies is to provide a global understanding of the physics of the folding transition in proteins [3]. Since the underlying energy landscape in proteins is believed to be rugged [4], it seems reasonable to suggest that some of the ideas developed for polymer models in which the monomers interact via quenched random variables may, in fact, be useful in understanding "phase transitions" in proteins. It is in this spirit that an assortment of models (to be collectively referred to as random heteropolymer models) have been studied [1–4]. So far, the only approach used in the study of the random heteropolymer models and other disordered polymer problems has been the replica method combined with a variational treatment [1,2,5]. These approximations, although widely used, have been shown to be exact only in the limit of large dimensionality for the problem of random manifolds [5(f)]. More importantly, the utility of the replica variational method has been questioned for random heteropolymer models in general [6]. Although the replica variational calculations have yielded a number of interesting results, their general validity has not yet been fully established.

In order to shed light on this difficult and important class of disordered Hamiltonian systems, we introduce a dynamical treatment. In this Letter, we study the dynamics of the random hydrophobic-hydrophilic model suggested recently by Garel, Leibler, and Orland (GLO) [7] which turns out to be simpler than the original random bond heteropolymer model. The random site model is perhaps a better caricature of proteins than the random bond models [1,2] because of the clear distinction between hydrophobic and hydrophilic residues. In addition, from a technical point of view, GLO have shown using replica methods that there is no replica symmetry breaking in the equilibrium treatment. The dynamical theory, which we treat using the same type of approximation as the static calculation, is meant to complement the equilibrium treatment of GLO [7]. In the process, new aspects of this class of models emerge (see below). We provide a self-consistent one loop solution of the dynamical equations and establish the following results which appear to be valid for a certain class of heteropolymer models.

(a) As in most models, the heteropolymer chain collapses from a high temperature random coil structure at a temperature T_Θ to a compact conformation. T_Θ is essentially determined by the effective strength of hydrophobic interactions.

(b) For the same set of parameters (i.e., when the hydrophobic interactions dominate) there is *a length scale dependent freezing temperature* $T_G(l)$. Depending on the ratio of the effective hydrophobic interaction strength to the disorder strength, $T_G(l)$ can be greater or less than T_Θ. When $T_G(l)$ is less than T_Θ the scale dependent freezing temperature offers a natural connection with global features of protein folding scenarios.

The GLO model [7] consists of a polymer chain with N monomers and is described by the set of position vectors at its nodes. It is convenient to work with continuum models with the positions in the chain parametrized by the contour length s such that $O \le s \le N$. We adopt the usual Edwards model. Since the heteropolymer is in solution, interactions induced by solvents have to be taken into account. This corresponds to the fact that a polymer made up of hydrophobic (hydrophilic) residues tend to collapse (swell) in solutions, and hence, when the chain is in solution, an effective attractive (repulsive) interaction is generated. The solvent induced interaction at each site is assumed to be random. The random interaction is assumed to depend only on the site s, and its strength is denoted by $\lambda(s)$. In accordance with the aforementioned discussions the GLO Hamiltonian is

$$\beta\mathcal{H} = \frac{3}{2}\int_0^N ds\left(\frac{dc}{ds}\right)^2 + \frac{1}{2}\int_0^N ds\int_0^N ds' \times \{v_0 + \beta[\lambda(s) + \lambda(s')]\}\delta[c(s) - c(s')] + \beta\mathcal{H}_1\,, \tag{1a}$$

where

$$\beta\mathcal{H}_1 = \frac{\omega_3}{3!}\int_0^N ds\int_0^N ds'\int_0^N ds''\,\delta[c(s)-c(s'')]\delta[c(s'')-c(s')] + \frac{\omega_4}{4!}\int_0^N ds\int_0^N ds'\int_0^N ds''\int_0^N ds'''\,\delta[c(s)-c(s'')]\delta[c(s'')-c(s')]\delta[c(s')-c(s''')], \quad (1b)$$

and $c(s)$ is the position vector of residue at point s. The random site variables are assume to have the Gaussian distribution

$$P[\lambda(s)] = \Omega\exp\left[-\frac{1}{2\lambda^2}\int_0^N ds\,[\lambda(s)-\lambda_0]^2\right]. \quad (2)$$

In the work of GLO the replica calculations leading to the mean field equations were studied using variational techniques. Here, using relaxational dynamics (thus avoiding replicas), we arrive at a picture of freezing transition which does not appear to have been noticed within the equilibrium statistical mechanical calculation [7].

We assume a Langevin equation for the relaxational dynamics

$$\frac{1}{\Gamma_0}\frac{\partial\vec{c}(s,t)}{\partial t} = -\frac{\partial(\beta\mathcal{H})}{\delta\vec{c}(s,t)} + \xi(s,t), \quad (3a)$$

where Γ_0 is a kinetic coefficient, and $\xi(s,t)$ is a Gaussian random force with zero mean and correlation given by

$$\langle\xi_\alpha(s,t)\xi_\beta(s',t')\rangle = \frac{2}{\Gamma_0}\delta_{\alpha\beta}\delta(s-s')\delta(t-t'). \quad (3b)$$

By use of standard field-theory techniques [8,9], the generating functional averaged over $\lambda(s)$ can be written as

$$[Z] = \int Dc\int D\hat{c}\exp[L_0 + V_2 + W_3 + W_4 + D], \quad (4a)$$

where L_0 is the free part of the action

$$L_0(c,\hat{c}) = \int ds\int dt\left\{i\hat{c}(s,t)\left[-\frac{1}{\Gamma_0}\frac{\partial c(s,t)}{\partial t} - \frac{3}{2}\frac{\partial^2 c}{\partial s^2} + \frac{i}{\Gamma_0}\hat{c}(s,t)\right] + i\hat{c}(s,t)\hat{l}(s,t) + c(s,t)l(s,t)\right\}, \quad (4b)$$

and the W_3 and W_4 terms arise from three and four body interactions in $\beta\mathcal{H}_1$, and V_2 is

$$V_2 = \frac{(V_0 + 2\beta\lambda_0)}{2}\int_0^N ds\int_0^N ds'\int dt\, i[\hat{c}(s,t)-\hat{c}(s',t)]\cdot\nabla\delta[c(s,t)-c(s',t)]. \quad (5)$$

The relevant disorder average term that leads to the scale dependent freezing is D, which can be written as

$$D[\hat{c},c] = \int\frac{d^3p}{(2\pi)^3}\int\frac{d^3q}{(2\pi)^3}\int dt_1\int dt_2\,\Delta(p,t_1,q,t_2), \quad (6a)$$

with

$$\Delta(p,t_1,q,t_2) = \mu\int ds\int ds_1\int ds_2[ip\cdot\delta\hat{c}(s,s_1,t_1)][iq\cdot\delta\hat{c}(s,s_2,t_2)]\exp\{i[p\cdot\delta c(s,s_2,t_2) - q\cdot\delta c(s,s_2,t_1)]\}, \quad (6b)$$

where $\mu = \beta^2\lambda^2/2$ and

$$\delta c(s,s_1,t_1) = c(s,t_1) - c(s_1,t_1),$$
$$\delta\hat{c}(s,s_1,t) = \hat{c}(s,t_1) - \hat{c}(s_1,t_1). \quad (6c)$$

From Eq. (5) it follows immediately that when the two body term vanishes the heteropolymer chain undergoes a collapse transition to compact structure at

$$T_\Theta = -2\lambda_0/v_0k_B. \quad (7)$$

Note that λ_0 is negative for an effective hydrophobic interaction. To investigate the possibility of freezing transitions we study whether correlations between various segments of the chain persist over time. In order to investigate this possibility a suitable order parameter is required, and we introduce this by defining the correlation function,

$$Q_{ss_1s_2}(t_1-t_2) = \langle\delta c(s,s_1,t_1)\cdot\delta c(s,s_2,t_2)\rangle, \quad (8)$$

where the average $\langle\cdots\rangle$ is over the action given in Eq. (5a). An order parameter, reminiscent of the Edwards-Anderson order parameter [10], can be defined as

$$q_{ss_1s_2} = \lim_{|t_1-t_2|\to\infty} Q_{ss_1s_2}(t_1-t_2). \quad (9)$$

This definition of the order parameter (as opposed to one in terms of monomer density variables) is extremely useful once we identify δc as the natural variable in the disorder averaged action. In order to make this order parameter similar to the ones that are encountered in the usual spin glass theories, we introduce the restriction $s_1 \approx s_2$, i.e., $|s_1 - s_2|/N \ll 1$. With this restriction, one can interpret the order parameter in Eq. (9) as the relative displacement of two points on the chain (s and s_1) projected approximately onto itself (s and s_2) after a long time. A nonzero value is clearly indicative of ordering

on a certain scale. The proximity of s_1 and s_2 leads to the following approximate relation for the equal time correlation function, namely,

$$\langle|\delta c(s,s_1t)|^2\rangle = l^2(s,s_1) \approx l^2(s,s_2)\,, \tag{10a}$$

$$Q_{ss''s'}(0) \approx l^2(s,s'')\,, \tag{10b}$$

where $l(s,s'')$ is the physical distance between s and s''.

In order to obtain a dynamical equation for the correlation function we follow the technique of Kirkpatrick and Thirumalai (KT) [11] of introducing the half-Fourier transform

$$\hat{Q}_{ss''s'}(\omega) = \int_0^\infty Q_{ss''s'}(t)e^{i\omega t}\,dt \tag{11}$$

which leads, in the ergodic phase, to the relation

$$\hat{Q}_{ss''s'}(\omega) = \frac{Q_{ss''s'}(0)}{-i\omega + \Gamma_{ss''s'(\omega)}/Q_{ss''s'}(0)}\,, \tag{12}$$

where $\Gamma_{ss''s'}(\omega)$ is a renormalized kinetic coefficient. The calculation of $\Gamma_{ss''s'}(\omega)$ is done in the Hartree approximation which is similar in spirit to the recent work of Cugliandolo, Kurchan, and P. Le Doussal [12]. From Eqs. (4) and (6) we note that there is a contribution to this order to D which has the structure $\delta\hat{c}\delta\hat{c}$, and this coefficient renormalizes the Γ_0 occurring in L_0. This approximation, together with the methods described by KT [11], allows us to obtain the renormalized $\Gamma(\omega)$ which, in a self-consistent one loop calculation, can be written as (details to be published elsewhere)

$$\Gamma^{-1}_{ss_1s_2}(\omega) = \Gamma_0^{-1} + \frac{\mu}{6}\int\frac{d^3p}{(2\pi)^3}\int\frac{d^3q}{(2\pi)^3}\left(\frac{(p\cdot q)^2}{3}\right) \times e^{-[p^2l^2(s,s_1)/6+q^2l^2(s,s_2)/6]}\cdot\hat{Q}_{ss_1s_2}(\omega)\,. \tag{13}$$

The above equation, which is obtained by expanding Eq. (9) in powers of δc, is valid to order $Q^2_{ss_1s_2}$. If we define $\phi(t) = Q_{ss_1s_2}(t)/Q_{ss_1s_2}(0)$, it follows that $\phi(t)$ satisfies in the time domain the following integrodifferential equation:

$$\frac{1}{\Gamma_0}\,\partial_t\phi(t) + \phi(t) + \overline{\mu}\int_0^t dt_1\phi(t-t_1)\partial_{t_1}\phi(t_1) = 0\,, \tag{14a}$$

where

$$\overline{\mu} = Q_{ss_1s_2}(0)\frac{\mu}{6}\int\frac{d^3p}{(2\pi)^3}\int\frac{d^3q}{(2\pi)^3}\,\frac{(p\cdot q)^2}{3} \times\exp\left[-\frac{1}{6}\,(p^2l^2 + q^2l'^2)\right]. \tag{14b}$$

If we assume that $Q_{ss_1s_2}(0)$ is continuous at the freezing temperature, then Eqs. (14) predict that a nonergodic transition to a frozen phase takes place at a characteristic temperature. This transition is identified by letting $t\to\infty$ and seeking a physical solution for $q_{\rm EA}$ where $q_{\rm EA} \equiv \lim_{t\to\infty} Q_{ss_1s_2}(t)$. We find that $q_{\rm EA}$ satisfies

$$\frac{q_{\rm EA}}{Q(0)} = \overline{\mu}q_{\rm EA}[1 - q_{\rm EA}/Q(0)]\,. \tag{15}$$

From Eq. (14b), $\overline{\mu}$ can be calculated as

$$\overline{\mu} = \frac{\mu}{[Q(0)]^4}\left(\frac{324}{512\pi^3}\right) = \frac{\lambda^2}{2(k_BT)^2}\left(\frac{324}{512\pi^3}\right)\frac{1}{[Q(0)]^4}\,. \tag{16}$$

It follows from Eq. (15) that a nonzero $q_{\rm EA}$ corresponding to a freezing transition is possible if

$$\lambda^2/k_BT^2 > \left(\frac{256\pi^3}{81}\right)Q^3(0) = \left(\frac{256\pi^3}{81}\right)\frac{1}{l^6}$$

or

$$k_BT < \alpha\lambda/l^3 \equiv k_BT_g(l)\,, \tag{17}$$

where $\alpha = (81/256\pi^3)^{1/2}$. In obtaining Eq. (17), which introduces the scale l, we have used $Q(0) = l^2$ [see Eq. (10b)]. The above equation predicts the existence of a scale dependent freezing as the temperature is lowered. This implies that the short length scale fluctuations freeze first, while the longer length fluctuations freeze at much lower temperatures. This is the principal result of our work.

The nonzero freezing temperature is a consequence of the randomness of the hydrophobic-hydrophilic interaction whose strength is given by λ. If $\lambda = 0$, then $T_g(l) \equiv 0$ [see Eq. (17)]. Consequently, the scale dependent glass transition temperature does not arise in homopolymers whose globular conformation at $T < T_\theta$ is determined by a balance between two body and three body interactions [see Eq. (1)].

There are a few implications of this work that we briefly outline here.

(1) Consider the freezing temperature (or the glass transition temperature) at a scale on the order of the entire chain R. This temperature is given by $k_BT_G(R) = \alpha\lambda/R^3$. We have two temperature scales, T_Θ [cf. Eq. (7)] and $T_G(R)$, in the problem and we can anticipate potentially three possible scenarios: (i) $T_\Theta < T_G(R)$; in this case disorder effects dominate to such a great extent that the glassy behavior manifests itself in the regime when the chain is still in the coil state. Here there is no possible ordering on at least the scale of the entire chain. (ii) $T_\Theta \geq T_G(R)$; here the chain collapses first (i.e., at $T < T_\Theta, R$ scales as $N^{1/d}$), and, as the strength of disorder is changed, $T_G(R)$ satisfying the condition $T_\Theta \geq T_G(R)$ becomes possible. General arguments based on caricatures of proteins suggest that this is the scenario most appropriate for the folding of random sequences of proteins. (iii) $T_\Theta \gg T_G(R)$; this would correspond to a small disorder limit. Here there is a whole range of temperatures over which glassy effects are not predominant even though the chain is in a collapsed phase. We speculate that this could, in the case of proteins, correspond to "equilibrium" molten globule phase [13], where, in spite

of the compact structure, there can be large conformational entropy.

(2) It is interesting to use the temperature scales T_Θ and $T_G(R)$ to compute their ratios as a function of the number of monomers, N. At $T < T_G$ the chain is stabilized by the three body term [see Eq. (16)], and the size of the polymer is given by

$$R^3 = \frac{\omega_3 N T T_\theta}{|\lambda_0|(T_\theta - T)}. \tag{18}$$

If we combine Eqs. (7), (18), and (19) we get

$$\frac{T_G(R)}{T_\Theta} = 1 + \left(\frac{4\lambda_0}{\lambda}\right)\frac{\omega_3 N}{\alpha v_0^2}\left(\frac{T_G(R)}{T_\Theta}\right)^2. \tag{19}$$

In the strong disorder limit (i.e., $|\lambda_0| \approx \lambda$) the constant $4\lambda_0/\alpha\lambda \approx 10$. Letting $\beta = \omega^3/v_0^2$, the equation for $T_\Theta/T_G(R)$ becomes

$$\frac{T_\Theta}{T_G(R)} = \frac{20\beta N}{\sqrt{1 + 40\beta N} - 1}. \tag{20}$$

One of the characteristics of proteins is that at a temperature T_F (typically less than T_Θ) the polypeptide chain makes a transition to a unique ground state with a well defined topology. It has been argued elsewhere that generically $\max(T_F) = T_\Theta$ [14]. This implies that the maximum value of T_F/T_G is given by the right hand side of Eq. (20). If the three body interaction is weak, i.e., $\beta \ll 1$, then it follows that $\max(T_F/T_G) = 1$. For $\beta \gg 1$ (not achievable in practice) the ratio T_F/T_G can be made arbitrarily large. Realistic values of β are in the range $0 < \beta \ll 1$. The typical values of β (which are proportional to the strength of the three body interaction) lie in the range $10^{-4}-10^{-3}$. In this range of β values, $T_F/T_G(R)$ ranges from 1.1 to 2 as N varies from 27 to 200 which are the typical numbers of aminoacid residues in naturally occurring proteins. It has been postulated that foldable proteins should have large values of T_F/T_G [15]. For this model at least an upper bound for this ratio for typical values of N found in proteins is in accordance with this expectation.

(3) The connection between the dynamical theory presented here and the static calculation of GLO [7] for this class of models is not entirely clear. However, there is some similarity between the nature of the transition discussed here and in p-spin glass models [11] and the associated spherical version [16]. It is clear that the trial function used in the static treatment by GLO [7] ignores the possibility of the scale dependence of $T_g(l)$. Another reason for the emergence of $T_g(l)$, that is apparently absent in the static treatment, is the following: The relaxation of the chain to equilibrium occurs on a time scale that diverges as the dimensionality and the length of the chain increase. This would naturally give rise to a difference in $T_g(l)$ and a static glass transition temperature. This possibility has been recognized in other disordered systems [17].

In summary, we have provided a dynamical treatment of the random hydrophobic-hydrophilic copolymer model and have shown that there is a freezing transition temperature (not to be confused with the folding temperature in proteins) that depends on the scale. Larger scale structures freeze at lower temperatures than shorter scale structures. Calculations similar to that done here for random bond models seem to suggest that this phenomenon could be generally applicable for random heteropolymers [18]. The scale dependent freezing temperature is a manifestation of topological frustration.

We are grateful to T. R. Kirkpatrick for several useful and interesting discussions. We thank an anonymous referee for insightful remarks and for bringing Refs. [12] and [17] to our attention. This work was supported in part by a grant from the National Science Foundation (through Grant No. CHE 96-29845).

[1] T. Garel and H. Orland, Europhys. Lett. **6**, 307 (1988).
[2] T. Garel and H. Orland, Europhys. Lett. **6**, 597 (1988).
[3] See, for example, J. D. Bryngelson and P. G. Wolynes, Proc. Natl. Acad. Sci. **84**, 7524 (1987); J. D. Bryngelson and P. G. Wolynes, J. Phys. Chem. **93**, 6902 (1989).
[4] P. G. Wolynes, J. Onuchic, and D. Thirumalai, Science **267**, 1619 (1995).
[5] (a) S. F. Edwards and M. Muthukumar, J. Chem. Phys. **89**, 2435 (1988); (b) E. Shakhnovich and A. Gutin, Biophys. Chem. **34**, 187 (1989); (c) J. D. Honeycutt and D. Thirumalai, J. Chem. Phys. **90**, 4542 (1989); (d) M. Sasai and P. G. Wolynes, Phys. Rev. Lett. **65**, 2740 (1990); (e) M. Mezard and G. Parisi, J. Phys. I (France) **1**, 809 (1991); (f) M. Mezard and G. Parisi, J. Phys. I (France) **2**, 2231 (1992); (g) P. Goldbart and A. Zippelius, Phys. Rev. Lett. **71**, 2256 (1993).
[6] S. Franz, M. Mezard, and G. Parisi, Int. J. Neural Syst. Suppl. **3**, 195 (1992).
[7] T. Garel, L. Leibler, and H. Orland, J. Phys. II (France) **4**, 2139 (1994).
[8] C. DeDominicis, Phys. Rev. B **18**, 49 131 (1978).
[9] H. Sompolinski, Phys. Rev. Lett. **47**, 359 (1981); H. Sompolinski and A. Zippelius, Phys. Rev. B **25**, 6860 (1982).
[10] See M. Mezard, G. Parisi, and M. Virasaro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[11] T. R. Kirkpatrick and D. Thirumalai, Phys. Rev. Lett. **58**, 2091 (1987); Phys. Rev. B **36**, 5388 (1987).
[12] L. F. Cugliandolo, J. Kurchan, and P. Le Doussal, Phys. Rev. Lett. **76**, 2390 (1996).
[13] O. B. Ptitsyn, J. Protein Chem. **6**, 273 (1987).
[14] D. Thirumalai, J. Phys. I (France) **5**, 1457 (1995).
[15] R. A. Goldstein, Z. A. Luthey-Schulten, and P. G. Wolynes, Proc. Natl. Acad. Sci. **89**, 4918 (1992).
[16] A. Crisanti and H.-J. Sommers, Z. Phys. B **8**, 341 (1992).
[17] L. F. Cugliandolo and D. S. Dean, J. Phys. A **28**, 4213 (1995).
[18] D. Thirumalai and J. K. Bhattacharjee (to be published).