PHYSICAL REVIEW E VOLUME 61, NUMBER 2 FEBRUARY 2000

Glassy dynamics of protein folding

Erkan Tüzel[1] and Ayşe Erzan[1,2]
[1]*Department of Physics, Faculty of Sciences and Letters, Istanbul Technical University, Maslak 80626, Istanbul, Turkey*
[2]*Feza Gürsey Institute, P. O. Box 6, Çengelköy 81220, Istanbul, Turkey*
(Received 10 September 1999)

A coarse-grained model of a random polypeptide chain, with only discrete torsional degrees of freedom and Hookean springs connecting pairs of hydrophobic residues is shown to display stretched exponential relaxation under Metropolis dynamics at low temperatures with the exponent $\beta \simeq 1/4$, in agreement with the best experimental results. The time dependent correlation functions for fluctuations about the native state, computed in the Gaussian approximation for real proteins, have also been found to have the same functional form. Our results indicate that the energy landscape exhibits universal features over a very large range of energies and is relatively independent of the specific dynamics.

PACS number(s): 87.17.Aa, 05.70.Ln, 64.70.Pf, 82.20.Rp

A huge amount of effort has recently been invested in modeling the interactions responsible for yielding the native states of proteins as their thermodynamic equilibrium state [1,2]. It has recently begun to be appreciated that such features of real proteins as the density of vibrational energy states [3] may be reproduced by coarse-grained model Hamiltonians that capture the essential mechanism driving the folding process, namely, hydrophobic interactions [3–9]. In this paper we introduce and study a model of N coupled, overdamped torsional degrees of freedom with discrete allowed states. Under Metropolis Monte Carlo dynamics, with random initial conditions, we find that at low temperatures the model exhibits power law relaxation for the initial stages of decay, and at the later stages the relaxation obeys a stretched exponential with the exponent $\beta \simeq 1/4$. This type of relaxation behavior is of the Kohlrausch-Williams-Watts type as observed experimentally for real proteins [1,10–12]. We find that at zero temperature the probability distribution function of the energy steps encountered along a relaxation path in phase space also obeys a stretched exponential form, with another exponent $\alpha \simeq 0.39$. We show that $\beta = \alpha/(\alpha+1)$, yielding a value for β, which is in very good agreement with our simulation results.

We take as our point of departure the model proposed by Haliloğlu, Bahar, and Erman [6]. The central idea of this model is that all interactions in the protein are governed by confining square-law potentials, so all attractions may be treated as if the residues interact with each other through Hookean forces [6–8]. This feature mimicks the effective pressure that is exerted on the hydrophobic residues of a randomly coiled molecule by the ambient water molecules.

To keep our model very simple, we consider covalent bonds as fixed rods of equal length. The residues located at the vertices may be polar P or hydrophobic H. All the hydrophobic vertices are to be connected to each other with springs of equal stiffness. This results in their being driven to the relatively less exposed center of the molecule in the low lying energy states, whereas the polar residues are closer to the surface (see Fig. 1), a feature that is common to the native configurations. Since the square law potential is an effective attraction induced by hydrophobic interactions with the surrounding water molecules, there is no contradiction in the growth of the restoring force with the distance, once the hydrophobic instability has set in. However, due to the absence of any cutoff in the range of the effective interaction, the initial stages of the dynamics may well be unrealistic and should be regarded with caution. Finally, the constraints placed on the conformations due to the rigid chemical bond lengths and restriction of the chemical and dihedral angles to discrete values prevent the molecule from disintegrating or collapsing to a point.

It is known that real proteins are distinguished by H-P sequences that lead to unique ground states while a randomly

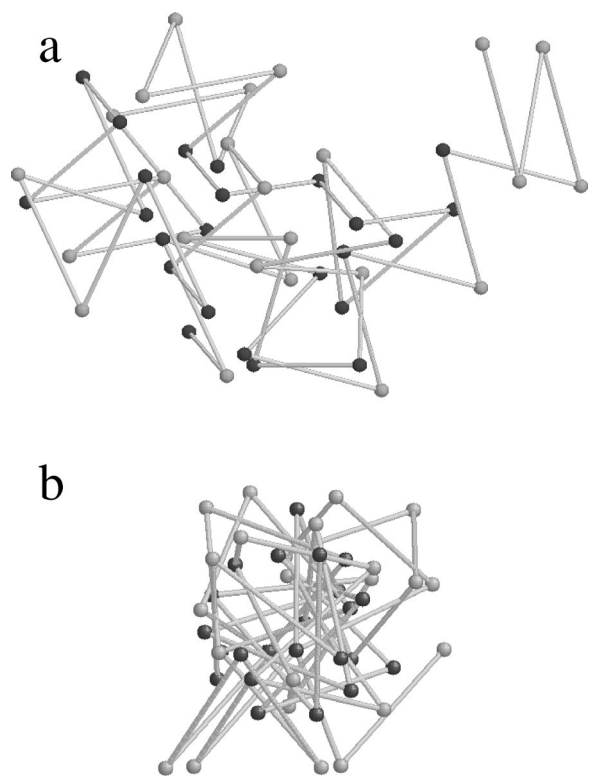

FIG. 1. Chain of $N=48$ residues, half of which are randomly chosen to be hydrophobic, (darker beads) shown (a) in a random initial configuration and (b) in a folded state reached under Metropolis dynamics. The chain has folded in such a way as to leave the polar residues on the outside. (Generated using RasMol V2.6.)

 ©2000 The American Physical Society

chosen H-P sequence will typically give rise to a highly degenerate ground state. Nevertheless, in our Monte Carlo study we considered a generic H-P sequence obtained by choosing 50% of the residues to be hydrophobic and distributing them randomly along the chain. In the absence of detailed knowledge regarding the rules singling out the realistic H-P sequences we believed this to be in keeping with our statistical approach. It might be speculated that the choice of equal probabilities for encountering H and P groups along the chain, and distributing them randomly, maximizes the configurational entropy of the chain [13] and enhances the ''designability,'' giving rise to rather realistic results.

Our model for the protein chain consists of N ''residues'' which are treated as point vertices, connected to each other by rigid rods. The ''bond angle'' α_i at the ith vertex, $i=1,\ldots,N-1$, is fixed to be $(-1)^i\alpha$, with $\alpha=68°$. The dihedral angles ϕ_i can take on the values of 0 and $\pm 2\pi/3$. The state (conformation) of the system is uniquely specified once the numbers $\{\phi_i\}$ are given. Thus, the residues effectively reside on the vertices of a tetrahedral lattice.

The energy of the molecule is

$$E=\frac{K}{2}\sum_{i,j} c_{i,j}|\mathbf{r}_i-\mathbf{r}_j|^2=K\sum_{i,j}\mathbf{r}_i^{\dagger}V_{ij}\mathbf{r}_j. \tag{1}$$

If we define $Q_i=1$ for the ith vertex being occupied by a hydrophobic residue, and $Q_i=0$ otherwise, we may write $c_{i,j}=Q_iQ_j$ and

$$V_{ij}=[(N_H-1)c_{i,i}-c_{i,j-1}-c_{i,j+1}]\delta_{i,j} -(1-\delta_{i,j})(1-\delta_{i,j-1}-\delta_{i,j+1})c_{i,j}. \tag{2}$$

The position vectors $\mathbf{r}_i$ of each of the vertices in the chain can be expressed in terms of a sum over the directors $\mathbf{R}_i$ of unit length representing the chemical bonds, which may be obtained from $\mathbf{R}_1$ by successive rotations $\mathbf{M}_k(\alpha_k)$ and $\mathbf{T}_k(\phi_k)$ through the bond and the dihedral angles [14],

$$\mathbf{r}_i=\sum_{j=1}^{i-1}\prod_{k=j}^{2}\mathbf{T}_k(\phi_k)\mathbf{M}_k(\alpha_k)\mathbf{R}_1, \tag{3}$$

where we may choose $\mathbf{R}_1$ along any Cartesian direction in our laboratory frame without loss of generality.

In order to investigate the relaxation properties of the present model, we have employed Metropolis Monte Carlo dynamics. This consisted of (a) choosing a pair (i,i') of dihedral angles randomly on the chain, and incrementing the $(\phi_i,\phi_{i'})$ in opposite directions by $\Delta\phi=\pm 2\pi/3$, (b) accepting the move with unit probability if $\Delta E\leqslant 0$ and with probability $p=\exp(-\gamma\Delta E)$ for $\Delta E>0$, (c) checking for an alternative path by repeating the second step once, before returning to the first step. Omission of this repeated trial would only change the scale of γ, which serves as an effective temperature. Since the system is strongly dissipative, strict (angular) momentum conservation is not crucial here. We monitor the relaxation of the total energy as a function of ''time'' measured in the number of MC steps, [i.e., the number of pairs (i,i') sampled] until a steady state is reached, typically in about 10 000 steps. The results for chains of $N=100$ averaged over 20 randomly chosen initial configurations at zero temperature ($\gamma=\infty$) are shown in Fig. 2. Defining $\epsilon\equiv(E-E_0)/E_I$, where E_0 is the (time-averaged) equilibrium energy and E_I the initial value, we find that it obeys a power law $\epsilon(t)\sim t^{-\sigma}$, with $\sigma=0.49\pm 0.01$ for the initial stages of the decay, while later stages can be fitted by a stretched exponential $\epsilon(t)\sim e^{-t^{\beta}}$, with $\beta=0.234\pm 0.003$. We also performed simulations for different values of γ, for chains of $N=48$, averaging over 100 runs with random initial configurations (see Fig. 3). For $\gamma\to\infty$, $\gamma=0.5$, and $\gamma=0.3$, the above relaxation behavior continues to hold and the exponents do not seem to depend on γ, with $\beta\simeq 1/4$ and $\sigma\simeq 1/2$, as given in Table I.

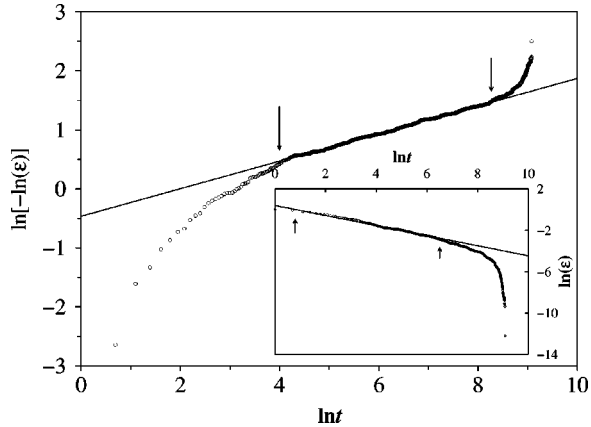

FIG. 2. Decay with time (in MC steps) of the energy [in units of $K/(\text{bond length})^2$] of an $N=100$ chain from a random initial configuration, along a zero temperature Metropolis trajectory, of 10 000 steps, averaged over 20 runs. The late stages fit on a stretched exponential curve $\epsilon(t)\sim\exp(-t^{\beta})$ with $\beta=0.234\pm 0.003$. The initial stage (inset) is fit by a power law $\epsilon(t)\sim t^{-\sigma}$ with $\sigma=0.49\pm 0.01$.

The variation of the total energy in time is sketched in Fig. 4 over a short sequence of relaxation events. Clearly one may write $E(t)$, averaged over many independent runs, as $\langle E(t)\rangle=\langle E(0)-\Sigma_{i=1}^{M}\Delta E_i\Theta(t-t_i)\rangle$ where Θ is the Heavyside step function and $t_i=\Sigma_{k=0}^{i-1}\tau_k$. Taking the time derivative one gets

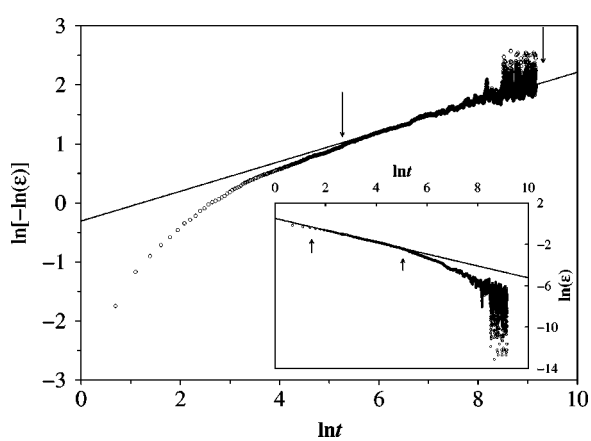

FIG. 3. Decay of the energy of an $N=48$ chain, along a Metropolis trajectory, from a random initial configuration averaged over 100 runs for $\gamma=0.3$. The initial stage (inset) is fit by a power law $\epsilon(t)\sim t^{-\sigma}$ with $\sigma=0.57\pm 0.01$, and the late stage to a stretched exponential with $\beta=0.25\pm 0.03$.

TABLE I. The exponent σ and β found for the power law and stretched exponential decay of the total energy with time, for different chain lengths N and inverse temperatures γ. The fits were obtained from a weighted least-squares computation.

N	γ	σ	$\Delta\sigma$	β	$\Delta\beta$
48	∞	0.57	0.01	0.281	0.004
	0.5	0.56	0.01	0.30	0.04
	0.3	0.57	0.01	0.25	0.03
100	∞	0.49	0.01	0.234	0.003

$$\langle \dot{E}(t)\rangle = \left\langle -\sum_{i=1}^{M} \Delta E_i \delta\left(t - \sum_{k=0}^{i-1} \tau_k \right) \right\rangle. \tag{4}$$

At zero temperature, the expectation value of $\dot{E}(t)$ can be calculated by carrying out an integration over the distibution of waiting times $\{\tau_k\}$, and the distribution of energy steps encountered along the relaxation path. The expectation value, $\langle \dot{E}(t)\rangle$ is then,

$$\langle \dot{E}(t)\rangle = -\left\langle \sum_{j=1}^{M} \Delta E_j \delta\left(t - \sum_{k=0}^{i-1} \tau_k \right) \right\rangle_{\Delta E, \tau}. \tag{5}$$

The distribution of waiting times τ_k is dependent only on the configuration of the chain at the kth step and independent of the previous waiting times. Since the dynamics is just changing a pair of dihedral angles in opposite directions, for each conformation $\{\phi_i\}$ one may define an associated chain of $N(N-1)/2$ sites, with each site corresponding to a pair (i,i') on the original chain. On the associated chain, a site will be assigned the value 1 if the corresponding pair has at least one ''allowed'' move, and the value 0 if both moves are ''blocked.'' Now the probabilities of encountering allowed or blocked moves as one implements the Metropolis dynamics outlined above are simply given by the density of 1's or 0's on the associated chain at a given relaxation step, namely, p_k and $q_k = 1 - p_k$. Therefore, in the kth conformation, the probability of making a transition after τ_k blocked moves simply obeys the first passage time distribution [15],

$$P_k(\tau_k) = \mu_k e^{-\mu_k \tau_k}, \quad \mu_k \equiv |\ln q_k|. \tag{6}$$

Writing the δ function in Eq. (5) in the Fourier representation and performing the τ-integrals we get

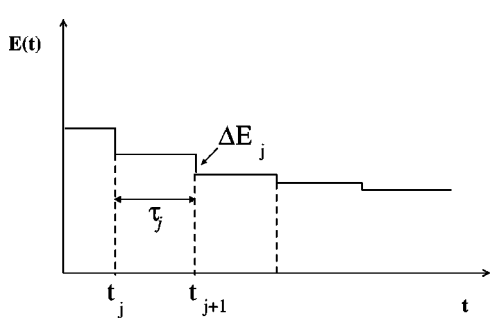

FIG. 4. Schematic plot of the variation of the total energy with time.

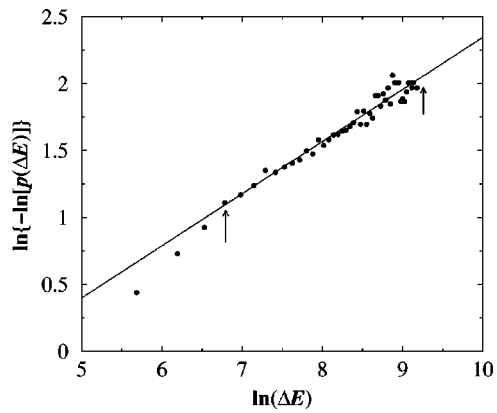

FIG. 5. Distribution of energy differences encountered along the relaxation path are fit to a stretched exponential. Level spacing histograms were formed for chains of $N=48$ and averaged over 100 runs for the zero-temperature Metropolis relaxation. The exponent α of the stretched exponential is found to be 0.39 ± 0.02.

$$\langle \dot{E}(t)\rangle = \frac{1}{2\pi} \sum_{j=1}^{M} \left\langle \Delta E_j \sum_{\ell=1}^{j-1} \prod_{\substack{k=0 \\ k\neq \ell}}^{j-1} \left(\frac{\mu_k}{\mu_k - \mu_\ell} \right) e^{-\mu_\ell t} \right\rangle_{\Delta E}. \tag{7}$$

We may argue that the larger the energy loss in a relaxation event, the longer it will take for the phase point to make a transition out of this state. Since μ_k is roughly the expectation for τ_k, we assume that $\mu_k \sim 1/\Delta E_k$. With the assumption that the energy steps encountered along a relaxation path are independently distributed, i.e., $P(\Delta E_1 \ldots \Delta E_M) = \prod_{s=1}^{M} P(\Delta E_s)$ for a process of M steps, one finds,

$$\langle \dot{E}(t)\rangle = -\frac{1}{2\pi} \sum_{j=1}^{M} \langle \Delta E_j \rangle \sum_{\ell=1}^{j-1} I_{j\ell}(t), \tag{8}$$

where $I_{j,\ell}(t)$ is

$$I_{j,\ell}(t) \equiv \int_0^\infty d(\Delta E_\ell) e^{-(t/\Delta E_\ell)} P(\Delta E_\ell) \times \left[\prod_{\substack{k=0 \\ k\neq \ell}}^{j-1} \left\langle \frac{\Delta E_\ell}{\Delta E_\ell - \Delta E_k} \right\rangle_{\Delta E_k} \right]. \tag{9}$$

We have determined from our simulations that the distribution $P(\Delta E_\ell)$ (see Fig. 5) also has the stretched exponential form $P(\Delta E_\ell) = P_o e^{-(\Delta E_\ell)^\alpha}$. The angular brackets then take the form

$$\Delta E_\ell \int_0^\infty (\Delta E_\ell - \Delta E_k)^{-1} \exp[-(\Delta E_k)^\alpha] d\Delta E_k \tag{10}$$

which we approximate by $\Delta E_\ell \exp[-(\Delta E_\ell)^\alpha]$. The integration in Eq. (9) is then straightforward, leading, upon substitution in Eq. (8), to

$$E(t) \sim t \sum_{j=1}^{M} \left(\frac{j-1}{j} \right) \exp(-a_j t^\beta), \tag{11}$$

where $a_j = j(1-\alpha)(\alpha j)^{-\beta}(1+\beta)^{-1}$ and

$$\beta = \frac{\alpha}{\alpha+1}. \tag{12}$$

Substituting the value of α we find from our simulations, namely, $\alpha = 0.39 \pm 0.02$, we get $\beta = 0.28 \pm 0.01$, which is the result we obtained from the fits to the MC simulations within our error bounds.

A study of the correlations between fluctuations about the native state [16] for the Beads-and-Springs model, with the *H-P* sequence and the contact map for the native states for seven real proteins (6lyz, 1cd8, 1bet, 1fil, 1bab, 1csq, and 1hiv) was performed by Erman [17]. Using the Gaussian approximation [18] to the coherent scattering function and a normal mode analysis [5,6,19,20], he also finds a stretched exponential relaxation with $\beta = 1/4$. Experiments on real proteins and polymers [1,10–12] yield $0.2 \leqslant \beta \leqslant 0.4$. Our results seem to be closer to 1/4 and smaller than the value most commonly found for spin-glasses [21], namely 1/3. It should also be noted that glassy behavior is obtained here in the absence of quenched randomness, or of frustration arising from steric hindrances, which we do not take into account.

Comparing the relaxation behavior near the native state with the behavior we observe at relatively high energies for random heteropolymers, we conclude that the relaxation behavior, and therefore the dynamics and the structure of the energy landscape, are universal over a very large range of energies, and are relatively independent of the specific sequence or the details of the dynamics.

We are grateful to Burak Erman for motivating this work and for the close interest he has shown in every stage of its progress and to Mustansir Barma for a useful discussion. One of us (A.E.) acknowledges partial support by the Turkish Academy of Sciences.

[1] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, Science **254**, 1598 (1991).
[2] P. G. Wolynes, J. N. Onuchic, and D. Thirumalai, Science **276**, 1619 (1995).
[3] D. ben-Avraham, Phys. Rev. B **47**, 14 559 (1993).
[4] K. A. Dill, S. Bromberg, K. Yue, K. M. Feibig, D. P. Yee, P. D. Thomas, and H. S. Chan, Protein Sci. **4**, 561 (1995).
[5] M. M. Tirion, Phys. Rev. Lett. **77**, 1905 (1996).
[6] T. Haliloğlu, I. Bahar, and B. Erman, Phys. Rev. Lett. **79**, 3090 (1997).
[7] B. Erman and K. Dill, J. Chem. Phys. (to be published).
[8] B. Erman (unpublished).
[9] E. Tüzel and A. Erzan, e-print cond-mat/9909350.
[10] J. L. Green, J. Fan, and C. A. Angell, J. Phys. Chem. **98**, 13 780 (1994).
[11] B. Erman and I. Bahar, Macromol. Symp. **133**, 33 (1998).
[12] J. Colmenero, A. Arbe, and A. Alegria, Phys. Rev. Lett. **71**, 2603 (1993).
[13] R. Mélin, H. Li, N. S. Wingreen, and C. Tang, J. Chem. Phys. **110**, 1252 (1999).
[14] P. J. Flory, *Statistical Mechanics of Chain Molecules* (Interscience, New York, 1969).
[15] W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1957), Vols. I and II.
[16] B. Erman, Comp. Theor. Polym. Sci. **9**, 295 (1999).
[17] B. Erman (private communication).
[18] B. Ewen and D. Richter, in *Elastomeric Polymer Networks*, edited by J. E. Mark and B. Erman (Prentice-Hall, Englewood Cliffs, NJ, 1992), p. 220.
[19] N. Go, T. Noguti, and T. Nishikawa, Proc. Natl. Acad. Sci. USA **80**, 3696 (1983).
[20] M. Levitt, C. Sander, and P. S. Stern, J. Mol. Biol. **181**, 423 (1985).
[21] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore 1987).