Configurational Diffusion on a Locally Connected Correlated Energy Landscape; Application to Finite, Random Heteropolymers

Jin Wang, Steven S. Plotkin and Peter G. Wolynes (*)

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

Abstract. — We study the time scale for diffusion on a correlated energy landscape using models based on the generalized random energy model (GREM) studied earlier in the context of spin glasses (Derrida B. and Gardner E., *J. Phys. C* **19** (1986) 2253) with kinetically local connections. The escape barrier and mean escape time are significantly reduced from the uncorrelated landscape (REM) values. Results for the mean escape time from a kinetic trap are obtained for two models approximating random heteropolymers in different regimes, with linear and bi-linear approximations to the configurational entropy *versus* similarity q with a given state. In both cases, a correlated landscape results in a shorter escape time from a meta-stable state than in the uncorrelated model (Bryngelson J.D. and Wolynes P.G., *J. Phys. Chem.* **93** (1989) 6902). Results are compared to simulations of the diffusion constant for 27-mers. In general, there is a second transition temperature above the thermodynamic glass temperature, at and above which kinetics becomes non-activated. In the special case of an entropy linear in q, there is no escape barrier for a model preserving ultrametricity. However, in real heteropolymers a barrier can result from the breaking of ultrametricity, as seen in our non-ultrametric model. The distribution of escape times for a model preserving microscopic ultrametricity is also obtained, and found to reduce to the uncorrelated landscape in well-defined limits.

1. Introduction

The study of biomolecular dynamics in general and protein folding in particular has inspired a good deal of study of the dynamics of disordered systems [1]. Statistically defined energy landscapes naturally emerge in biomolecular physics because of the heteropolymeric complexity of a biomolecule's sequence. Many questions about the energy landscapes and dynamics of biomolecules resemble the issues in the theory of ergodicity breaking in glasses and spin glasses [2] but the mesoscopic scale of protein molecules means that hopping between globally distinct minima, which in the thermodynamic limit would be strictly inaccessible, must be taken into account, and may indeed dominate the experimental behavior [1]. Starting with the work of Bryngelson and Wolynes [3] the connection of the dynamics on statistical energy landscapes and folding has been exploited both qualitatively to understand the perplexing complexity ·of

---

(*) Author for correspondence (e-mail: wolynes@aries.scs.uiuc.edu)

folding kinetics experiments and quantitatively in interpreting simulations [4,5] and developing prediction algorithms [6]. Foldable proteins contain an energetic bias towards a folded state which is likely due to selection by natural evolution. On the other hand, as in typical mean field spin glasses, there are many uncorrelated states which can act as kinetic traps, at least transiently. Dynamics in the glassy phase for these systems has been studied using random energy models [7,8]. The escape from these traps determines the effective diffusion coefficient for flow of the ensemble of protein structures towards the folded state needed to give fast folding on biologically relevant timescales. This general picture has been confirmed semiquantitatively through the use of lattice models of minimally frustrated heteropolymers [9,10]. The original BW analysis [3] assumed that the rugged parts of the energy landscape could be modeled in the most extreme form as a completely random landscape described by the random energy model of Derrida [11]. Superimposed on this was the bias due to minimal frustration. The random energy model belongs to a general class of disordered systems which lack special symmetries (e.g. time inversion symmetry in the Sherrington-Kirkpatrick model of a spin glass), and having a discrete jump in the order parameter $q(x)$ describing the replica symmetry breaking, such as that seen in Potts glasses [12]. Other aspects of the low energy states are universal, such as the statistics of their non-self-averaging processes below the transition, and some of this universality may carry over to the dynamics. The "first order" glass transition seen in random heteropolymers (RHPs) by replica methods [13] puts these systems in the same universality class as the REM as far as issues of ergodicity breaking are concerned.

    While the random energy model, being in the right universality class, describes in a renormalized sense the nature of a random heteropolymer (RHP) landscape, for a quantitative theory (see for example [14]), it is certainly relevant to take into account the correlations in the landscape which are inherent in the polymeric nature of the problem. Recently Plotkin, Wang and Wolynes have shown how the thermodynamics of glassy trapping [15] and folding [16] can be treated in this way. The thermodynamic glass transition temperatures are not much modified in many cases, although the nature of the replica symmetry breaking can change as a function of the density. While the relevant transition temperatures are well determined by the random energy model estimates, it was shown that the configurational entropy relevant to the number of basins of attraction was significantly modified. In the original BW analysis, the Levinthal paradox [17] arising from the consideration of the search time through possible conformational states on a flat energy landscape was shown to be connected on a globally rough landscape with the difficulties of escape from the traps on the random energy surface. Since the configurational entropy of the basins is changed in the correlated model, it is relevant to ask whether similar modifications occur when the kinetics are examined. In this paper we discuss the kinetics of escape from traps on a correlated energy landscape for two models approximating random heteropolymers (RHP). The first (hereafter referred to as the REM strata model, RS model, or RSM) considers strata of states at a given similarity to a specified state, which are uncorrelated to each other, but indeed correlated in energy to the specific state. In the second model (hereafter the ultrametric model), an ultrametric hierarchy is imposed on all of the states (equilibrium or not), the statistics of which are analyzed in the context of the generalized random energy model (GREM) [18]. This microscopic ultrametricity is somewhat different than the conventional thermodynamic ultrametricity used to describe equilibrium states in mean field spin-glasses. The regions of validity for both of these models as determined by the accuracy of their description of the true organization of states of a random heteropolymer are an important issue. We analyze the dynamics for both models to obtain the average escape time from a metastable trap. In calculating this diffusion time, linear and bi-linear approximations to the configurational entropy as a function of similarity $S_c(q)$ are employed [15]. These simple forms are chosen primarily for their illustrative value; more complex forms such as those used
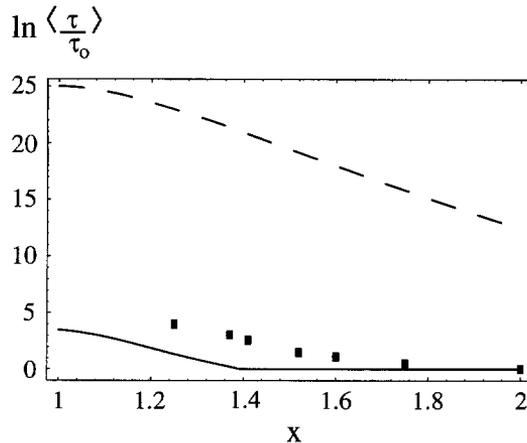
Fig. 1. — Logarithm of average diffusion time for the correlated landscape model of a 27-mer RHP, *versus* reduced temperature $x = T/T_G^0$. Solid line: the correlated RSM, with linear configurational entropy, and $S_0 = 27 \times 0.9$ [15]. Dashed line: results of $\langle \tau(x) \rangle$ for the REM (uncorrelated) model [3]. The data points are adapted from measurements of the diffusion constant from simulations on a specific 3-letter sequence 27-mer [5] (see text).

by us in describing the thermodynamics [16] can be accommodated *via* numerical work. We also neglect density variations, which for minimally frustrated polymers such as proteins, play a significant role in obtaining accurate barrier heights. However, for our analysis of generic undesigned random heteropolymers, the constant density approximation is most likely as accurate as the other approximations we have made (*e.g.* linear and bilinear approximations to the entropy). Incorporating an additional density variable is certainly possible in extending this analysis. Also, it has been shown previously that the glass temperature is lower than the collapse temperature in the thermodynamic limit [19], and it is in this regime that we are interested in the application of the GREM.

In the REM stratum model, and the ultrametric model with the special exception of the purely linear entropy case, there is a temperature T_A, higher than the glass temperature T_G, above which activated kinetics ceases. This is an explicit example of the behavior expected in a random first order phase transition [12,20]. For the linear $S_c(q)$ in the ultrametric model $T_A = T_G$.

We show in addition that correlations improve the agreement of the statistical landscape theory with simulations of the configurational diffusion constant, at least for small heteropolymer systems on a lattice (see Fig. 1).

The correlations are taken into account in this analysis only in a pairwise fashion, which means that some correlation effects involving local and separable motions of different parts of the chain are still not completely accurately treated in this model. These require a more elaborate analysis of the local reaction coordinates for escape from traps. Nevertheless, for moderate sized systems this analysis should show the general direction of the effects, and is the simplest generalization of the random energy model, beyond globally connected kinetic models [23]. In this paper we will concentrate primarily on the random heteropolymer problem. En passant we indicate the effect of minimal frustration but save for elsewhere the detailed discussion of the combined effects of local correlation and bias in minimally frustrated landscapes.

We should also mention that the approach used in this paper, though arrived at in the context of finite heteropolymers, is applicable as an approximation to any system describable by a locally connected, correlated energy landscape.

We organize this paper as follows: in Section 2, we review the description of the correlations in the landscape that are required to model the heteropolymer, and obtain the free energy relative to a given state. Free energy functions are obtained in two approximations, which in the following sections are used to calculate escape barriers from basins of attraction. In Section 3, we analyze the ground states and glass transition temperatures in the two models described in Section 2. In Section 4, we study the kinetics of escape processes for a small random heteropolymer (RHP) using a linear approximation of entropy as a function of specific contacts, which can be applied to lattice model 27-mers often used to mimic the protein folding process. In Section 5, we calculate the mean escape time from a basin of attraction for large RHP's using a piece-wise linear approximation to the entropy.

2. Correlated Free Energy Landscapes

In this section, we review the description of the energy statistics of a random heteropolymer (RHP) by a correlated energy landscape, highlighting those quantities used in the kinetic analysis.

For an RHP, each pair of interacting monomers mn has an interaction energy ϵ_{mn} that can be taken to be a Gaussian random variable. The contact Hamiltonian $\mathcal{H}$ for a given configuration is given by:

$$\mathcal{H} = \sum_{m<n} \epsilon_{mn}\sigma_{mn} \tag{2.1}$$

where $\sigma_{mn} = 1$ when there is a contact made between monomers mn in the chain, and $\sigma_{mn} = 0$ otherwise. Here contact means that the two monomers mn are within a small distance (contact radius) of each other. We assume as in lattice models that there is an additional part of the Hamiltonian which keeps the chain connected, and represents hard constraints, i.e. all connected configurations which are allowed have the same energy at least as far as this part of the Hamiltonian is concerned. This is used for lattice models, and is likely a good approximation for real proteins. In addition we assume the model contains hard excluded volume forces, which again do not bias different allowed states. Both chain connectivity and excluded volume contribute to the entropy, i.e. in the GREM approach we can separate the energetic and entropic issues, taking chain connectivity and excluded volume into account through the counting of states of a lattice chain as a function of imposed constraints, $S(q)$ (see Eq. (2.5)). There are potentially interesting purely dynamical effects connected with chain connectivity and excluded volume, however these are likely to be small for real, finite size proteins because so much of the protein actually lies on the surface.

Because of the small number of nearest neighbors for realistic lattice models of proteins, many investigations have shown the move set in the configurational search does not greatly alter the result, and the problem of self avoidance is not a primary one [9]. However the dynamical glass temperature T_A described later in this paper is expected to increase as a result of hard-core interactions. Indeed the hard core fluid shows a transition to activated behavior that depends on density alone [21].

Since the total energy of the polymer is a sum of many random variables (contact energies), it is approximately a Gaussian random variable with probability distribution

$$P(E) = \frac{1}{(2\pi\Delta E^2)^{1/2}} \exp[-\frac{E^2}{2\Delta E^2}], \tag{2.2}$$

with variance in energy $\Delta E^2 = N z_N \epsilon^2$, where z_N is the number of contacts per monomer, and ϵ is the width of the Gaussian energy distribution of a single contact (which sets the roughness energy scale of the random energy model). If we pick two different states of the polymer, and ask for the joint probability of state i having energy E_i and state j having energy E_j, the answer will depend on the two energy values chosen and also on how similar the states i and j are. Because the energy is a sum of pair terms, the appropriate measure of similarity is the overlap parameter q, defined as the fraction of contacts the two states i and j have in common

$$q = \frac{1}{N z_N} \sum_{m < n} \sigma_{mn}^i \sigma_{mn}^j, \tag{2.3}$$

where $\sum_{m<n} \sigma_{mn}^i = \sum_{m<n} \sigma_{mn}^j = N z_N$.

Given two configurations $\{\sigma_{mn}^i\}$ and $\{\sigma_{mn}^j\}$, the probability that they have energies E_i and E_j respectively is given by

$$\left\langle \delta \left[E_i - \mathcal{H} \left(\{\sigma_{mn}^i\} \right) \right] \delta \left[E_j - \mathcal{H} \left(\{\sigma_{mn}^j\} \right) \right] \right\rangle,$$

where $\langle .. \rangle$ means an average over the bond energy distributions. Performing this average gives the same joint probability distribution as in the GREM [18]

$$\frac{P(E_i, E_j | q)}{P(E_i)} \sim \exp \left(-\frac{(E_i - q E_j)^2}{2 \Delta E^2 (1 - q^2)} \right) \tag{2.4}$$

so that applying the thermodynamics of the GREM to the RHP with Gaussian interaction distributions gives an accurate measure of the free energy up to pair correlations of the partition function $\langle Z(T) Z(T') \rangle$.

By choosing a particular state i with energy E_i, we can study the thermodynamics of the system relative to this state ($E_j = E$). We eventually wish to find the escape time from state i via a barrier crossing process. To this end we obtain the free energy relative to this state, parameterized by the similarity q. A free energy $F(q)$ as a function of q (see Fig. 4 for some possible curves) implies a collection of states at any q, since many states can have overlap q with i. This is the idea behind the transition state ensemble used for complete folding of minimally frustrated systems (see [16] and references therein). So there are many possible kinetic paths of escape from i, the number being determined by the configurational entropy $S_c(q)$ at similarity q. In a calculation of the $S_c(q)$ [15] (summarized briefly in Appendix A), we found that for small polymers, the entropy follows an approximately linear form, i.e.

$$S_c(q) = S_0 (1 - q). \tag{2.5}$$

For larger polymers, the configurational entropy is approximated by a bi-linear (piece-wise linear) form, i.e.

$$\begin{aligned} S_c(q) &= S_0 - \alpha q \quad \text{for} \quad q < q_0 \\ S_c(q) &= \gamma(1 - q) \quad \text{for} \quad q > q_0 \end{aligned} \tag{2.6}$$

with $|\gamma| < |\alpha|$, or equivalently

$$\begin{aligned} S_< &= S^\dagger - \frac{(S_0 - S^\dagger)}{q^\dagger} (q - q^\dagger) \qquad q < q^\dagger \\ S_> &= S^\dagger - \frac{S^\dagger}{1 - q^\dagger} (q - q^\dagger) \qquad q > q^\dagger \end{aligned} \tag{2.7}$$

with $S^\dagger < S_0 \left(1 - q^\dagger\right)$. We will use either of these notations depending on which one is more advantageous in elucidating the results. Here $q_0 = q^\dagger$ is the approximate crossover point where the entropic behavior changes from bond formation to sequence melting [15] (see Appendix A).

The energies of the states at q are distributed according to equation (2.4), so that the microcanonical entropy is given by [25]

$$S(E, q, E_i) \cong \ln \left(\exp \left[S_{\mathrm{c}}(q)\right] \frac{P\left(E, E_i | q\right)}{P\left(E_i\right)} \right). \tag{2.8}$$

In this model there is no explicit overlap q' between states in the stratum at q, so that their correlations with each other are not considered. The mean and width of the Gaussian distribution of energies for the states at q are explicitly q dependent due to their correlations with state i, but in writing (2.8) we assume the relative fluctuations in $n(E, q, E_i) = \exp S(E, q, E_i)$ are negligible, and $n(E, q, E_i)$ may be replaced by its average in the argument of the log (hence the REM stratum model). This assumption is still reasonable when the number of states is large, but for smaller $n(E, q, E_i)$, it becomes important to investigate the distribution $P\left(n\left(E, q, E_i\right)\right)$, or at least moments beyond the first, e.g. $\langle n(E_1, q, E_i) n(E_2, q, E_i)\rangle$. Another point to mention regarding (2.8) is that the description allows for the existence of "non-ultrametric" states having less overlap with each other than they do with state i. This description can be said to be accurate when uncorrelated states (with $q_{12} = 0$ for states 1 and 2) predominate the number distribution.

Proceeding within the REM strata model, we use the thermodynamic condition $1/T = \partial S(E, q, E_i)/\partial E$ to obtain the thermal energy and the thermal entropy [26]

$$\begin{aligned} E\left(q, T\right) &= q E_i - \frac{\Delta E^2}{T}\left(1 - q^2\right) \\ S\left(q, T\right) &= S_{\mathrm{c}}(q) - \frac{\Delta E^2}{2T^2}\left(1 - q^2\right) \end{aligned} \tag{2.9}$$

from which the free energy of the band of states having overlap q with a given one can now be written as

$$F_{\mathrm{rs}}\left(q, T\right) = q E_i - T S_{\mathrm{c}}(q) - \frac{\Delta E^2}{2T}\left(1 - q^2\right) \quad \text{for } T > T_{\mathrm{G}}\left(q\right), \tag{2.10}$$

where $T_{\mathrm{G}}(q)$ is given in equation (3.3). A free energy of this form was used to describe the barrier in the folding transition [16].

The free energy obtained this way has in it the *a priori* assertion that there exists a state i with energy E_i, which induces or designs the aspect of minimal frustration into the heteropolymer. This, coupled with the fact that there can exist states with overlaps $< q$ at each stratum (non-ultrametric states), allows there to be states with lower energy than the REM ground state $E_{\mathrm{GS}}^0 = -\sqrt{2 S_0 \Delta E^2}$. This comes about because the ground state of the collection of states having overlap q with i ($E_i = E_{\mathrm{GS}}^0$) is actually lower than E_{GS}^0, due to the fact that the correlated average of the distribution approaches E_{GS}^0 faster than its width decreases due to the REM stratum approximation. Because of this the barrier comes from deep states that are correlated to i, but only weakly correlated to each other. Clearly the barrier in this case must be interpreted as escape from a basin of attraction rather than a specific state. The problem of minimizing the number of these states which are correlated but distinct from a designed state is related to the problem of negative design. This effect of a correlated landscape was recently investigated in the context of optimized Hamiltonian approaches to structure prediction [27].

For strata of highly similar states, it becomes increasingly likely that states in this band are correlated to each other as well. In this regime, an ultrametric structure is a better

approximation to the organization of states of the heteropolymer. To this end we approximate the free energy of the heteropolymer using the GREM hierarchy, which attributes contributions to the energies of states on each branch of the ultrametric tree (see Ref. [18] and Appendix B). This model has a well-defined ground state energy, and we expect that the escape barriers, which in the RSM came from ultrametricity-breaking of meta-stable states, will be significantly modified.

Appendix B contains calculations of: 1) the distribution of escape times from a metastable state at temperature T, 2) The distribution of free energy functions relative to a state i with energy E_i, both calculated in the context of the GREM, with an ultrametric organization of states. The most probable form of the free energy as a function of q obtained there is

$$F_u(q,T) = qE_i - TS_c(q) - \frac{\Delta E^2}{2T}(1-q) \quad T > T_G^0. \tag{2.11}$$

3. Ground States and Glass Temperatures

Below T_g, the kinetics is non-self averaging in that the rate of escape from a kinetic trap depends not just on its energy but differs from one trap to another. More elaborate tools need to be developed in order to account for these non-self-averaging effects, which have applications to both molecular evolution and combinatorial synthesis. However we can make significant headway by investigating the energies at the boundaries of statistical distributions of states, and the temperatures at which the system is frozen into these energies.

The ground state of the system is the energy at which the number of states $n(E) = \exp(S_0) P(E) \approx \mathcal{O}(1)$, where S_0 is essentially the total number of states and $P(E)$ is given by equation (2.2):

$$E_{GS}^0 = \pm\sqrt{2S_0\Delta E^2}. \tag{3.1}$$

The system reaches the negative root of (3.1) at temperature $T_G^0 = \sqrt{\Delta E^2/(2S_0)}$.

Asserting the existence of a state i with energy E_i modifies the ground state energies. For the RS model, the ground state in each stratum q is the energy where the exponent of (2.8) = 0,

$$E_{GS}(q) = qE_i \pm \sqrt{2\Delta E^2(1-q^2)S_c(q)} \tag{3.2}$$

which the system reaches at the q dependent glass temperature

$$T_G(q) = \sqrt{\frac{(1-q^2)\Delta E^2}{2S_c(q)}}. \tag{3.3}$$

For illustration, we approximate the configurational entropy as a linear function of constraint q, $S_c(q) = S_0(1-q)$ (an accurate approximation for small polymers), and assume a designed ground state i with $E_i = E_{GS}^0$. The RS model predicts that the designed state freezes first at $\sqrt{2}T_G^0$, and then a gradual freezing occurs outward from the state i at $T_G^{linear} = \sqrt{(1+q)\Delta E^2/(2S_0)}$ to which the system is frozen into energy $E_{GS}^{linear}(q) = E_{GS}(q+(1-q)\sqrt{1+q})$, until a ground state at $q = 0$ of E_{GS}^0 is frozen in at T_G^0. The free energy at T_G^0 then has a minimum at $q \cong 0.48$ of $\cong 1.11E_{GS}^0$ and maxima at $q = 0$ and $q = 1$ of E_{GS}^0.

In an ultrametric model we can somewhat more rigorously investigate the freezing of pairs of states with overlap q. Let two states a and b have energies $E_a = \phi_< + \phi_>^a$ and $E_b = \phi_< + \phi_>^b$ (as in Fig. 5 with $\phi_>^i = \phi_>^a$ and $\phi_> = \phi_>^b$). The log number of pairs with overlap q is $S_{PAIRS} = S_0 + S_c(q)$. The ground state energy of $\phi_<$ at q, for two states having energies E_a

and E_b, is the value of $\phi_<$ where the log number of pairs

$$\ln\Omega_q\left(\phi_<, E_a, E_b\right) = S_0 + S_c\left(q\right) - \frac{1}{2\Delta E^2}\left(\frac{\phi_<^2}{q} + \frac{(E_a - \phi_<)^2}{1-q} + \frac{(E_b - \phi_<)^2}{1-q}\right)$$

vanishes. This gives

$$\phi_<^{\rm GS} = \frac{q\left(E_a + E_b\right)}{1+q} \pm \frac{q^{1/2}}{1+q}\left[2\left(1 - q^2\right)\left(S_0 + S_c\left(q\right)\right)\Delta E^2 - \left(E_a^2 - 2qE_aE_b + E_b^2\right)\right]^{1/2} \quad (3.4)$$

For a, b two ground states and a linear $S_c(q)$, $\phi_<^{\rm GS}$ in (3.4) is frozen into its most probable value (see Eq. (B.1))

$$\phi_<^{\rm GS}\left(q\right) = qE_{\rm GS}^0 \tag{3.5}$$

at a temperature given by

$$\frac{1}{T_{\rm G}^{\phi<}} = - \left.\frac{\partial\ln\Omega_q\left(\phi_<, E_a, E_b\right)}{\partial\phi_<}\right|_{\phi_<^{\rm GS}}$$

or

$$T_{\rm G}^{\phi<} = \sqrt{\frac{\Delta E^2}{2S_0}} = T_{\rm G}^0. \tag{3.6}$$

An analogous calculation gives $\phi_>^{\rm GS} = (1 - q)E_{\rm GS}^0$ at $T_{\rm G}^{\phi>} = \sqrt{\Delta E^2/(2S_0)}$. So for an ultrametric model with a linear $S_c(q)$, the freezing is q independent and equal to $T_{\rm G}^0$, as would be the case from a direct calculation in the GREM model [18]. The q-independent free energy for $T \leq T_{\rm G}^0$ is equal to $E_{\rm GS}^0$, so there are no states with lower energy than $E_{\rm GS}^0$. This will have consequences for the kinetic escape barrier above $T_{\rm G}^0$, discussed in the next section.

4. Barriers and Escape Times for Small Polymers

In this section we calculate the mean escape time $\langle\tau\left(T\right)\rangle$ from a basin of attraction for a small RHP, $i.e.$ the average time required for the RHP to reconfigure to unrelated states by an activated transition out of the basin of attraction. The inverse of this quantity is approximately equal to the diffusion constant on a rugged but correlated landscape.

For small polymers such as those describable by 27-mer lattice models, Plotkin, Wang, and Wolynes [15] have shown that the configurational entropy crudely approximates a linear function of constraint q (2.5). The mapping to lattice models should apply to significantly larger proteins than the sequence length of the corresponding lattice model, depending on the amount of secondary structure present [4].

For an ultrametric model, it is simple to see that for a purely linear configurational entropy there are no escape barriers above the thermodynamic glass temperature; barriers only appear below $T_{\rm G}$ when the system is in the glass phase. The free energy (2.11) is a linear function of q with slope

$$TS_0 + \frac{\Delta E^2}{2T} + E_i.$$

Now the largest escape barriers would occur when E_i happens to be a ground state, $E_i = E_{\rm GS}^0$. Letting $T = xT_{\rm G}^0$ where $x \geq 1$, this slope becomes $\propto (-2 + x + 1/x)$, which is ≥ 0 for all $x \geq 1$ (all $T \geq T_{\rm G}^0$). Then the maximum of the free energy F is at the state itself ($q = 1$) and escape is always a downhill diffusive process, with no activation barrier.

The situation is significantly different when the free energy is better approximated by REM strata (cf. Eq. (2.10)). For the purpose of obtaining escape times, we have set $\overline{E}(Q) = 0$ and $\Delta E^2(Q) = \Delta E^2$ to simplify the calculation. We first note that because of the gradual onset of freezing (see comments after Eq. (3.3)) starting at $\sqrt{2}T_G^0$ down to T_G^0, The free energy (2.10) is split into three temperature regimes:

$$T > \sqrt{2}T_G^0 \qquad F = F_m(q, E_i, T) = qE_i - TS_0(1-q) - \frac{\Delta E^2}{2T}(1 - q^2)$$

$$T_G^0 < T < \sqrt{2}T_G^0 \left\{ \begin{array}{ll} F = F_m(q, E_i, T) & q < q_G(T) \\ F = F_f(q, E_i) = qE_i - (1-q)\sqrt{2S_0\Delta E^2(1+q)} & q > q_G(T) \end{array} \right.$$

$$T < T_G^0 \qquad F = E_{GS} = -\sqrt{2S_0\Delta E^2}\left(q + (1-q)\sqrt{1+q}\right) \qquad (4.1)$$

where $q_G(T) = (2S_0T^2/\Delta E^2) - 1$ is the q value at temperature T where the frozen and unfrozen ("melted") free energy meet, i.e. the inverse of $T_G(q)$ (Eq. (3.3) with a linear $S_c(q)$).

We seek the diffusion constant on the rough landscape of an RHP, approximated as $D(T) \approx 1/\langle\tau(T)\rangle$, by calculating the mean escape time $\langle\tau(T)\rangle$ at temperature T:

$$\langle\tau(T)\rangle = \int_{E_{GS}^-}^{E_{GS}^+} dE_i P(E_i, T)\tau(E_i, T) \qquad (4.2)$$

where $P(E_i, T)$ is the thermal distribution of energies at temperature T (see Eq. (B.6)), and $\tau(E_i, T)$ is defined below. The structure of the free energy in this model is such that there exists a free energy minimum at a value of $q \equiv q^\ddagger$ between $q = 0$ and $q = 1$. The barrier to escape from a given state ($q = 1$) to unrelated states ($q = 0$) is then just $F(q = 0) - F(q^\ddagger)$. The escape time from a metastable state of energy E_i is essentially a Boltzmann factor of the activation energy, times the number of ways to cross the barrier:

$$\tau(E_i, T) = \tau_0 \exp\left(\frac{F(q = 0, T) - F(q^\ddagger, T, E_i)}{T}\right) \qquad (4.3)$$

where for proteins τ_0 is on the order of micro seconds [24]. The largest contributions to $\langle\tau(T)\rangle$ in (4.2) come from low energy states that have a barrier to escape. Higher energies that have no barrier (downhill) have escape times $\cong \tau_0$.

For purposes of calculation, we scale the temperature and energy to be in units of the REM glass temperature and REM ground state:

$$\begin{array}{lll} T &=& xT_G^0 \qquad\qquad x \geq 1 \\ E &=& yE_{GS}^0 = -y\sqrt{2s_0\Delta E^2} \quad 0 \leq y \leq 1. \end{array} \qquad (4.4)$$

First note that there are typically no macroscopic barriers above $\sqrt{2}T_G^0$. The argument is as follows: for typical thermal energies $E_i = -\Delta E^2/T$, the high temperature free energy (4.1) is $\propto (-x - 1/x) + q(x - 2/x) + q^2/x$. There is no macroscopic barrier when the free energy minimum $\partial F/\partial q = 0$ occurs at $q^\ddagger = 0$ or $x = \sqrt{2}$. So above $T = \sqrt{2}T_G^0$, kinetics is typically non-activated.

A more detailed calculation in Appendix C using equation (4.2) achieves the same result, plus the result that precisely at $T = \sqrt{2}T_G^0$, the mean escape time scales as a power law in system size ($\langle\tau(\sqrt{2}T_G^0)\rangle \sim \mathcal{O}(N^{1/2})$). We mention again that in calculating the mean escape time,

the integral over the region having an activation barrier dominates, and this region determines the limits of the integral.

For temperatures $T_G^0 < T < \sqrt{2}T_G^0$ where the system is partially frozen, the free energy (4.1) is split into two regimes: melted below $q_G(x) = x^2 - 1$ and frozen above $q_G(x)$. Even within this temperature range the calculation of the barrier and corresponding escape time must be split again into two temperature regions. The higher temperature region is defined by $q_G(x) > q^\ddagger$ ($y = 1$), which means that the minimum of the free energy is always in the melted region (since $q^\ddagger(y) = x(2y - x)/2 < q^\ddagger(1)$), and the barrier is always calculated from F_m in equation (4.1). Then the form of the expression for the escape time is identical to that above $\sqrt{2}T_G^0$, but the expressions in the exponent are now > 0, resulting in barriers that are now exponentially large below $\sqrt{2}T_G^0$. The crossover temperature $x^\# T_G^0$ occurs when $q_G(x^\#) = q^\ddagger(y = 1, x^\#)$ or $x^\# = (1 + \sqrt{7})/3$. Below this temperature of $\cong 1.22\,T_G^0$, there are energies where the escape barrier must be calculated between the frozen minimum $F_f(q = q^\ddagger)$, and the melted maximum $F_m(q = 0)$. Expressions for the escape time in this temperature regime are split into two terms, one contribution from high energies and one from low energies. As the temperature approaches T_G^0 the mean escape time becomes

$$\left\langle \tau\left(T_G^0\right)\right\rangle = \frac{\tau_0}{\sqrt{\pi S_0}} \exp\left(0.225 \times S_0\right) \tag{4.5}$$

and the activation energy (energetic barrier) to jump from trap to trap is

$$\Delta F\left(T_G^0\right) = \Delta E\left(T_G^0\right) \cong 0.1126\left|E_{GS}^0\right|. \tag{4.6}$$

The escape time here is significantly smaller than in the REM, where $\left\langle \tau\left(T_G^0\right)\right\rangle_{rem} = \tau_0 \exp S_0$, although in both cases the escape barrier scales exponentially with system size N. For a 27-mer, $\left\langle \tau\left(T_G^0\right)\right\rangle \approx 2 \times 10^3 \tau_0$, whereas $\left\langle \tau\left(T_G^0\right)\right\rangle_{rem} \approx 10^{15}\tau_0$. A correlated landscape is smoother, which results in a reduced search time. A consequence of this is that the number T_F/T_G, which characterizes the degree of minimal frustration in a heteropolymer, does not need to be so large for natural proteins as would appear from previous estimates based on the REM. The escape barrier at T_G^0 for a correlated landscape is also smaller than in the REM ($\Delta E_{rem} = T_G^0 S_0 = E_{GS}/2$).

Figure 1 shows the log of the average diffusion time in units of τ_0 *versus* reduced temperature $x = T/T_G^0$, using the parameter S_0 appropriate for a collapsed 27-mer ($S_0 = 0.9 \times 27$). For the correlated landscape the results of the calculation of Appendix C can be summarized as

$$\langle \tau(T)\rangle = \begin{cases} \tau_0 & \text{for } \sqrt{2}T_G < T \\[2ex] \begin{aligned}&\frac{\tau_0}{2\sqrt{\pi S_0}}\left(\frac{cT_G}{T} - 1\right)^{-1}\left[\exp 2S_0\left(\frac{(g_1 + 1)T_G}{T} - 1 - \frac{T_G^2}{T^2}\right)\right.\\ &\left.+ \left(\frac{2(c - T/T_G)}{2 - (T/T_G)^2} - 1\right)\exp S_0\left(4 - \frac{3T_G^2}{T^2} - \frac{5T^2}{4T_G^2}\right)\right]\end{aligned} & \text{for } T_G < T < \sqrt{2}T_G \\[3ex] \frac{\tau_0}{2\sqrt{\pi S_0}}\left(\frac{1}{c - 1}\right)\exp 2S_0\left(g_1 - 1\right) & \text{for } T \leq T_G \end{cases} \tag{4.7}$$

where $T_G \equiv T_G^0 = \sqrt{\Delta E^2/(2S_0)}$, and c and g_1 are defined in Appendix C and are numerically given by $c \cong 1.477$ and $g_1 \cong 1.113$. We mention again that we neglect the coupling of polymer density with temperature, which can play a significant role. Also plotted in Figure 1 are

the results from the uncorrelated random energy landscape [3]:

$$\langle \tau(T) \rangle_{\mathrm{rem}} = \begin{cases} \tau_0 \exp\left(\dfrac{2S_0 T_{\mathrm{G}}^2}{T^2}\right) & \text{for } 2T_{\mathrm{G}} < T \\[2mm] \tau_0 \exp\left[S_0 - 2\left(\dfrac{1}{T_{\mathrm{G}}} - \dfrac{1}{T}\right)^2 S_0 T_{\mathrm{G}}^2\right] & \text{for } T_{\mathrm{G}} < T < 2T_{\mathrm{G}} \\[2mm] \tau_0 \exp S_0 & \text{for } T \le T_{\mathrm{G}} \end{cases} \qquad (4.8)$$

and the results from simulations on the 27-mer [5] measuring the diffusion constant for a specific 3-letter sequence ($1/D$ is plotted here, where $D = \Delta Q^2/\tau_{\mathrm{corr}} \approx 1/\tau_{\mathrm{corr}}$ is the diffusion constant in Q-space). Our theory does not estimate the number τ_0 in the simulations, which depends on the details of the move set. Hence the values of the simulation data are normalized so that at high temperatures the measured correlation time coincides with τ_0, the barrier-less escape time. Notice that the escape time on the correlated landscape is much lower than the uncorrelated result. In addition there is a temperature T_{A} above which kinetics is non-activated, again differing from the REM. Escape times calculated for the correlated model by the analytic theory underestimate the escape times obtained in lattice simulations, but are much closer to the simulated values than the simple REM results.

Notice again that the diffusion time drops to τ_0 at $\sqrt{2}T_{\mathrm{G}}^0$, where the dynamics is no longer activated. Thus the correlated energy model possesses a transition which the REM does not, from a high temperature regime with essentially no escape barriers to a low temperature regime with escape barriers. This kind of behaviour with two characteristic temperatures (T_{g} and T_{A}) is expected for the Potts glass and some of the more sophisticated models of the random heteropolymer [12,20]. The behavior in the correlated energy model considered here is similar, but the transition is not sharp for finite N. There are always some activated transitions from traps.

For studying the special case of diffusion in a minimally frustrated polymer, we let Q represent the overlap with the native state, and q represent overlap among states with a given Q (there are still many conformational states with a given overlap Q with respect to the native state). In other words, folding progresses along the native order parameter by a typically activated diffusional rate $R(Q)$, which is evaluated by finding the escape time of meta-stable states at each stratum Q. In each stratum, the energy distribution of states is given by $P(E_i) \approx \exp\left[-\left(E_i - \overline{E(Q)}\right)^2/2\Delta E^2(Q)\right]$, and the configurational entropy is $S(Q)$. In general, states must have energies higher than $E_i^{c-} = \overline{E(Q)} - \sqrt{2\Delta E^2(Q)S(Q)}$ and lower than $E_i^{c+} = \overline{E(Q)} + \sqrt{2\Delta E^2(Q)S(Q)}$, where the statistical numbers are macroscopic.

To a first approximation, the folding time of a minimally frustrated heteropolymer can then be estimated as

$$\tau_{\mathrm{FOLD}} \approx \max_Q \left[\tau(Q) \exp\frac{\Delta F(Q)}{T}\right] \qquad (4.9)$$

where $\Delta F = F(Q) - F(Q_{\mathrm{unfold}})$ is the thermodynamic folding free energy barrier, and $\tau(Q)$ is the escape (diffusion) time in a particular stratum of similarity Q to the ground state. It is this quantity $\tau(Q)$ which is calculated below. The REM estimate for the escape time is given by equation (4.8). The calculations in the remainder of this paper improve on these estimates of the escape time (the exponentially large prefactor to the folding time due to transient trapping in Eq. (4.9)), by taking into account correlations in the energy landscape. The time $\tau(Q)$ is related to the diffusion constant D in the Kramers theory of folding [5] by $\tau(Q) \approx 1/D$,

to within a constant of order one determined by the mean squared change of the order parameter in a single escape event.

5. Barriers and Escape Times for Larger Heteropolymers

As mentioned in Section 2, the configurational entropy for larger polymers can be approximated by a bilinear form [15] (*cf.* Eqs. (2.6, 2.7)). We can proceed with this form of the entropy to find the average diffusion time at temperature T for larger heteropolymers. Here the effects of the form of the configurational entropy on the kinetic glass temperature can be readily seen, and in the appropriate limit the REM result can be re-obtained.

5.1. DIFFUSION AMONG ULTRAMETRICALLY ORGANIZED STATES. — The analysis is similar to that in the previous section with the simplification that the free energy is linear in q, however it is a piece-wise function of q. Using the scaling in equation (4.4), the piece-wise free energy relative to state i with energy E_i is

$$
\begin{aligned}
F_< &= \sqrt{\frac{S_0 \Delta E^2}{2}} \left[-x - \frac{1}{x} + q \left(\frac{1}{x} + \left(1 - \frac{S^\dagger}{S_0} \right) \frac{x}{q^\dagger} - 2y \right) \right] & q < q^\dagger \\
F_> &= \sqrt{\frac{S_0 \Delta E^2}{2}} \left[-\frac{S^\dagger}{S_0 \left(1 - q^\dagger \right)} x - \frac{1}{x} + q \left(\frac{1}{x} + \frac{S^\dagger}{S_0 \left(1 - q^\dagger \right)} x - 2y \right) \right] & q > q^\dagger.
\end{aligned} \tag{5.1}
$$

This piece-wise linear free energy has no escape barrier from i at $q = 1$ when the slope of $F_>$ is ≥ 0. This occurs naturally at high energies or $y \leq y_A (x)$ where

$$
y_A (x) = \frac{S^\dagger}{2 S_0 \left(1 - q^\dagger \right)} x + \frac{1}{2x} \tag{5.2}
$$

Equation (5.2) sets an upper limit to the energies contributing to the escape time.

Using the typical (thermal) energies at temperature T in (5.2), *i.e.* $E_i = -\Delta E^2 / T$ or $y_A = 1/\tilde{x}_A$, gives a crude estimate of the temperature $\tilde{T}_A$ above which escape barriers vanish and kinetics is non-activated. While this estimate will turn out not to give very accurate numerical values of the kinetic transition temperature T_A, it is useful in describing the trends in T_A in the limits of linear configurational entropy ($S^\dagger / S_0 = 1 - q^\dagger$) and a cratered, REM-like landscape ($S^\dagger, q^\dagger \to 0$). The results will be numerically crude because energies deeper than the thermal energy contribute most to the escape time.

Solving (5.2) for $\tilde{x}_A$ gives

$$
\tilde{x}_A = \frac{\tilde{T}_A}{T_G^0} = \sqrt{\frac{S_0}{S^\dagger} \left(1 - q^\dagger \right)}
$$

or

$$
\tilde{T}_A = \sqrt{\frac{\Delta E^2 \left(1 - q^\dagger \right)}{2 S^\dagger}}. \tag{5.3}
$$

Note $\tilde{T}_A \to \infty$ when $S^\dagger \to 0$. Losing all the entropy at $q^\dagger$ means that the system must jump out of a large energetic barrier before it gains entropy to compensate (but note in this approximation $q^\dagger$ needs not $\to 0$). This means that effectively we have a "cratered" free energy landscape, on which kinetics is always activated at all temperatures. This limit reproduces the result of the REM analysis, which also has such a pock-marked landscape because states are uncorrelated, and there is no "knowledge" of the existence of a low energy state until that state is reached.

When $S^\dagger/S_0 = 1 - q^\dagger$, $S_c(q)$ is a purely linear function of q. For this special case $\tilde{T}_A = T_G^0$ as obtained in the last section, and there are no activation barriers until $T \leq T_G^0$. Energetic losses are always exactly balanced by entropic gains. A more accurate calculation of T_A will be given below.

Equation (5.3) is precisely the higher temperature scale of the two step GREM for a given bilinear $S_c(q)$. [18] Following Derrida's analysis of the GREM, a bilinear $S_c(q)$ with $S^\dagger/S_0 < 1 - q^\dagger$, has only one thermodynamic glass temperature $T_G^0 = \sqrt{\Delta E^2/(2S_0)}$ above which the system is in its unfrozen phase. The above analysis gives significance to the higher temperature T_A, which did not enter into the thermodynamic analyses.

The barrier height (over T) to escape from a state i at temperature xT_G^0 is

$$\frac{\Delta F}{T} = S_0 \left[\frac{2\left(1 - q^\dagger\right)}{x} y - \frac{S^\dagger}{S_0} - \frac{\left(1 - q^\dagger\right)}{x^2} \right] \tag{5.4}$$

which vanishes for high energy states $E_i > y_A E_{GS}^0$, and is largest at T_G^0 ($x = 1$ and $y = 1$) when it is given by

$$\Delta F = \sqrt{\frac{S_0 \Delta E^2}{2}} \left(1 - q^\dagger - \frac{S^\dagger}{S_0} \right). \tag{5.5}$$

Note that (5.5) equals the REM barrier of $T_G^0 S_0 = E_{GS}^0/2$ when $S^\dagger, q^\dagger \to 0$ (again, a cratered landscape), and that when $S^\dagger/S_0 = 1 - q^\dagger$ (a linear $S_c(q)$), $\Delta F = 0$, as obtained in the last section.

The average escape time from equation (4.2) is then

$$
\begin{aligned}
\langle \tau(x) \rangle &= \tau_0 \sqrt{\frac{S_0}{\pi}} \exp{-S_0} \left(\frac{2}{x^2} + \frac{S^\dagger}{S_0} - \frac{q^\dagger}{x^2} \right) \times \int_{y_A(x)}^1 dy \, \exp S_0 \left[2y \frac{(2 - q^\dagger)}{x} - y^2 \right] \\
&= \frac{\tau_0}{2} \exp S_0 \left(\frac{2 - 3q^\dagger + (q^\dagger)^2}{x^2} - \frac{S^\dagger}{S_0} \right) \left\{ \mathrm{erf} \left[-\frac{\sqrt{S_0}}{x} \left(2 - q^\dagger - x \right) \right] \right. \\
&\qquad \left. - \mathrm{erf} \left[-\frac{\sqrt{S_0}}{2x} \left(3 - 2q^\dagger - \frac{S^\dagger}{S_0} \frac{x^2}{1 - q^\dagger} \right) \right] \right\}.
\end{aligned}
\tag{5.6}
$$

The integral is dominated by the ground state energy $y = 1$, i.e. the lowest energies contribute most to the average escape time. Using a large N approximation for the error functions, the mean escape (diffusion) time for a long RHP with ultrametric states is given by

$$
\langle \tau(T) \rangle \cong
\begin{cases}
\tau_0 & \text{for } T_A < T \\[2ex]
\dfrac{\tau_0}{2\sqrt{\pi S_0}} \left((2 - q^\dagger) \dfrac{T_G}{T} - 1 \right)^{-1} & \\
\quad \times \exp S_0 \left[2(2 - q^\dagger) \dfrac{T_G}{T} - (2 - q^\dagger) \dfrac{T_G^2}{T^2} - 1 - \dfrac{S^\dagger}{S_0} \right] & \text{for } T_G < T < T_A \\[3ex]
\dfrac{\tau_0}{2\sqrt{\pi S_0}} \left(1 - q^\dagger \right)^{-1} \exp S_0 \left(1 - q^\dagger - \dfrac{S^\dagger}{S_0} \right) & \text{for } T < T_G
\end{cases}
\tag{5.7}
$$

where $T_G \equiv T_G^0$. T_A, the temperature above which kinetics is typically non-activated, is given by the vanishing of the exponent in (5.7):

$$T_A \cong \frac{T_G^0}{1 + (S^\dagger/S_0)} \left(2 - q^\dagger + \sqrt{(2 - q^\dagger)\left(1 - q^\dagger - \frac{S^\dagger}{S_0} \right)} \right). \tag{5.8}$$
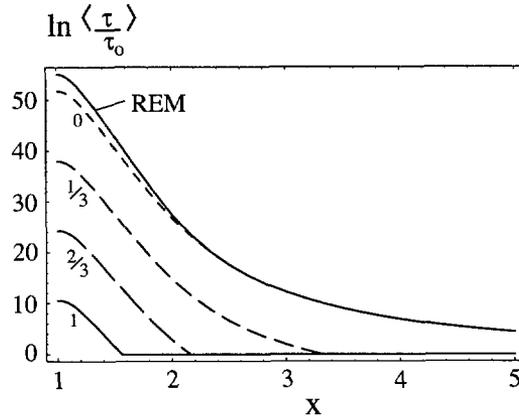
Fig. 2. — The escape time in the correlated model approaches the REM result in the limit of a cratered landscape. Plotted here is equation (5.6) with the parameters of a typical 64-mer : $S_0 \cong 0.9 \times 64$, $S^\dagger = 0.4 \times 64\alpha$ and $q^\dagger = 0.29\alpha$, where $\alpha = 1, 2/3, 1/3, 0$. The $\alpha = 0$ result essentially coincides with the REM result, with the small deviation coming from the prefactor.

The point $\tilde{x}_A$ given earlier ($\tilde{T}_A$ in Eq. (5.3)) is a point of inflection for the curve marking the onset of the decay of activation barriers for a finite system. However the escape time at these temperatures is still large. More accurate values of T_A can be obtained from (5.8), which give good agreement to the temperature where $\langle \tau(x) \rangle$ becomes $\cong \tau_0$. It can be seen from (5.8) that when the configurational entropy is purely linear ($S^\dagger/S_0 = 1 - q^\dagger$), $T_A \to T_G^0$, as obtained in the previous section. When $S_c(q)$ is linear ($S^\dagger/S_0 = 1 - q^\dagger$), the escape times are non-exponential and scale as $\sim N^{-1/2}$.

When $S^\dagger, q^\dagger \to 0$ the free energy landscape is highly cratered, and the escape time becomes

$$\langle \tau(T_G^0) \rangle \cong \frac{\tau_0}{2\sqrt{\pi S_0}} e^{S_0} \tag{5.9}$$

which reproduces the REM scaling with entropy. In other words, even though the landscape is correlated, the escape time agrees with the REM because the polymer must entirely reconfigure to escape a kinetic trap — there is little entropic gain until many bonds are broken. To obtain the behavior of T_A in the limit of a cratered landscape, we must use equation (5.6) rather than (5.7) since x_A is large and thus the error functions are small but not insignificant. For $S^\dagger, q^\dagger \to 0$, the value of y maximizing the integrand in (5.7) is $y_{\max} = 2/x$ (the energies with largest contribution become the statistically most probable energies for a cratered landscape), from which (5.7) becomes

$$\langle \tau(x) \rangle_{\mathrm{crat}} \sim \exp\left(\frac{2S_0}{x^2}\right), \tag{5.10}$$

from which we can see that kinetics is still activated as in the Ferry law, with exponentially large barriers, until $x \to \infty$ or $T_A \to \infty$.

The mean escape time from a basin of attraction, as given by equations (5.6, 5.7) is significantly less than the REM. In Figure 2, results are shown using equation (5.6) for the 64-mer with $S_{64}^\dagger \cong 0.4 \times 64$ and $q_{64}^\dagger \cong 0.29$. Results are also plotted as the landscape goes over to the REM (cratered) one, by taking $S^\dagger = \alpha S_{64}^\dagger$ and $q^\dagger = \alpha q_{64}^\dagger$ and letting $\alpha \to 0$.

The log of the mean escape time, using parameters appropriate for the 64-mer and 125-mer in equation (5.6), are shown in Figures 3a and 3b. The steepest descents approximation (5.7)
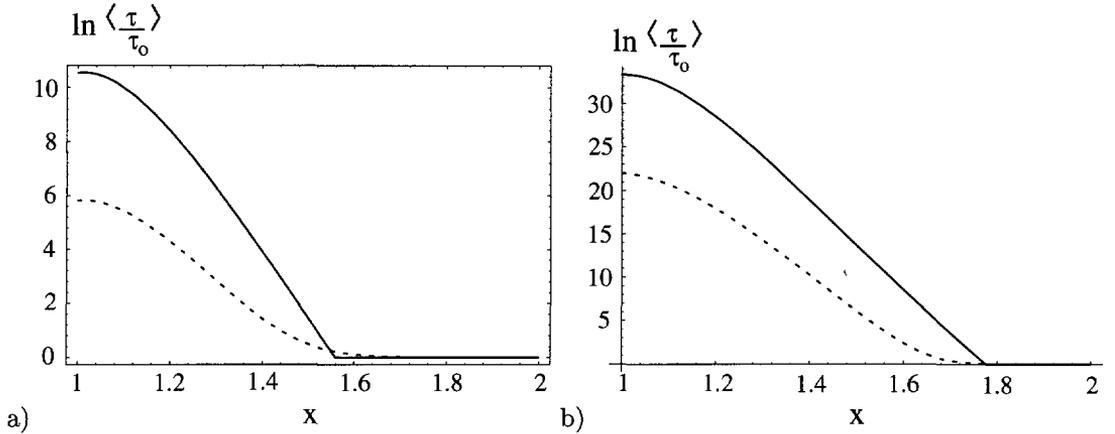
Fig. 3. — a) Logarithm of average diffusion (escape) time for the correlated landscape model of a 64 mer RHP *versus* reduced temperature $x = T/T_G^0$. A bi-linear approximation to the configurational entropy is used, with total entropy $S_0 \cong 0.9 \times 64$. Solid line: $\langle \tau(x) \rangle$ for the ultrametric model. Dashed line: $\langle \tau(x) \rangle$ for the REM stratum approximation. (The parameters used in the bilinear entropy formulae (2.6, 2.7) are $\alpha = 64 \times 1.55$, $\gamma = 64 \times 0.56$, or $S^\dagger = 0.4 \times 64$ and $q^\dagger = 0.29$.) b) Average escape time from a basin for a correlated landscape imitating a 125-mer. The curves are as in the 64-mer case. (The parameters used in the bilinear entropy formulae (2.6, 2.7) are $\alpha = 125 \times 2$, $\gamma = 125 \times 0.48$, $S^\dagger = 0.36 \times 125$ and $q^\dagger = 0.24$).

is nearly coincident with equation (5.6). Also plotted are the results of the RS model, described in the next section.

5.2. Diffusion in the Uncorrelated Stratum Approximation.

5.2. DIFFUSION IN THE UNCORRELATED STRATUM APPROXIMATION. — The methods and results of this section are somewhat involved, and the reader less interested in the detailed calculation may skip to the illustrative numerical results.

Now let us again examine the free energy expression as function of q, T and E_i in the uncorrelated stratum approximation (Eq. (2.10)). Knowing the free energy, we can calculate the rate for escaping from a given reference state of the energy E_i. To start, we search for the extrema of the free energy. The maximum of the free energy locates the barrier for the escape from a particular deep state we have chosen. Then the diffusion constant or mean life time can be calculated. The transition state position $q^\ddagger$, where the maximum occurs, defines the kinetic size of the basin of attraction for the given state. States inside this range ($q^\ddagger < q < 1$) have to overcome a free energy barrier to move to another basin. The new basin will be almost uncorrelated to the current one. We locate the barrier by taking the derivative of the free energy and setting it equal to zero. For a reference state with energy E_i, this gives:

$$E_i + \Delta E^2 q / T - T \mathrm{d}S/\mathrm{d}q = 0. \tag{5.11}$$

With the piece-wise linear entropy expression that mimics the random heteropolymer, this gives for $q < q^\dagger$:

$$q_{\min}^1 = -\frac{(E_i + T\alpha)T}{\Delta E^2} \tag{5.12}$$

and for $q > q^\dagger$:

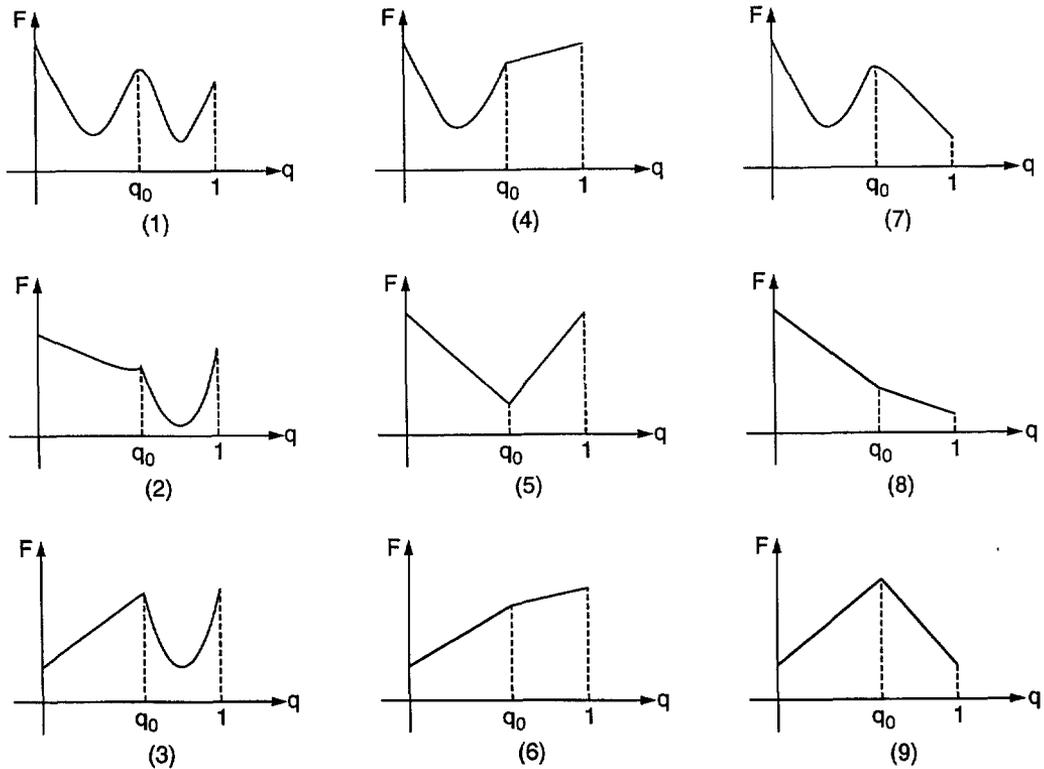$$q_{\min}^2 = -\frac{(E_i + T\gamma)T}{\Delta E^2}. \tag{5.13}$$

Fig. 4. — The 9 possible situations of the free energies *versus* overlap or fraction of contacts q are shown for the piece-wise linear approximation for the entropy.

It turns out that in these two regions the free energy may exhibit a minimum, but the only possible maximum will be at either $q^\dagger$ where the low entropy formula and high entropy formula meet or at two ends ($q = 0$ and $q = 1$). A variety of cases must be analyzed since the two minima at $q_{\min}^1$ and $q_{\min}^2$ behave differently in several different regimes. We give the detailed discussion in Appendix D.

We can classify the form of the free energy for all possible temperatures, reference energy and overlap regimes into 9 possible situations summarized in Figure 4. Detailed results for the various cases are shown in the Appendix D.

For the situation where the free energy in the high q regime is monotonically increasing and the free energy in the low q regime is either monotonically increasing (situation 6), monotonically decreasing (situation 5) or having a minimum (situation 4) of Appendix D, there is no barrier to escape away from the chosen state E_i. The general expressions of the parts of the thermally averaged escape time contributed by the given ranges of energies can be written as (assuming $\overline{E(Q)} = 0$ and $\Delta E(Q) = \Delta E$):

$$\tau'_{(4),(5),(6)} = \int \tau_0 p(E,T) = \frac{\tau_0}{2} \mathrm{erf}[\frac{\Delta E^2 + TX}{\sqrt{2}\Delta ET}]. \tag{5.14}$$

X is the integration limit of the energy E_i, given in the Appendix D.

For the situation where free energy in the high q regime is monotonically decreasing and free energy in the low q regime is either monotonically increasing (situation 9), monotonically decreasing (situation 8) or having a minimum (situation 7), there is indeed a barrier for escape.

For the situations (7) and (9) of Appendix D, the barrier is given by $F(q = q^\dagger) - F(q = 1)$, this contribution to the escape time from the appropriate energy ranges is $\tau'_{(7),(9)} = \int \tau_0 \exp[\frac{F(q=q^\dagger)-F(q=1)}{T}]p(E,T)$,

$$\tau'_{(7),(9)} = \frac{\tau_0}{2} \exp[\frac{\Delta E^2}{T^2}(1 - q^\dagger)^2 - \gamma(1 - q^\dagger)]\mathrm{erf}[\frac{\Delta E^2(2 - q^\dagger) + TX}{\sqrt{2}\Delta ET}] \qquad (5.15)$$

where X is the integration limit for the energy E_i, given in the Appendix D.

For situation (8) of Appendix D, the free energy barrier is given by $F(q = 0) - F(q = 1)$, the escape time contribution to the given energy range is $\tau'_{(8)} = \int \tau_0 \exp[\frac{F(q=q^\dagger)-F(q=1)}{T}]p(E,T)$. The result is:

$$\tau'_{(8)} = \frac{\tau_0}{2} \exp[\frac{\Delta E^2}{T^2} - S_0]\mathrm{erf}[\frac{2\Delta E^2 + X}{\sqrt{2}\Delta E}] \qquad (5.16)$$

where X is the integration limit for energy E_i, given in the Appendix D.

For the situation where the free energy in the high q regime gives a minimum while the free energy in the low q regime is either monotonically increasing (situation 3), monotonically decreasing (situation 2) or has a minimum (situation 1), there is again a barrier for escape from the vicinity of the minimum. This gives results similar to those discussed just above.

For situation (1) and (3) of the Appendix D, the free energy barrier is given by $F(q = q^\dagger) - F(q = q_{\min})$ where $q_{\min}$ is the minimum of $F(q)$ at $q > q^\dagger$. The escape time contribution to the given energy range is $\tau'_{(1),(3)} = \int \tau_0 \exp[\frac{F(q=q^\dagger)-F(q=q_{\min})}{T}]p(E,T)$. The result is:

$$\tau'_{(1),(3)} = \frac{\Delta ET\tau_0}{\sqrt{2\pi}(\gamma T^2 - \Delta E^2(1 - q^\dagger))} \exp[\gamma^2 T^2/(2\Delta E^2) + \gamma q^\dagger + (q^{\dagger^2} - 1)\Delta E^2/(2T^2)]$$
$$\times \exp[(\gamma T/\Delta E^2 + (q^\dagger/T - 1/T))X] \qquad (5.17)$$

where X is the integration limit for E_i, given in the Appendix D.

For situation (2), the free energy barrier is given by $F(q = 0) - F(q = q_{\min})$ where $q_{\min}$ is the minimum of $F(q)$ at $q > q^\dagger$. The escape time contribution to the given energy range is $\tau'_{(2)} = \int \exp[\frac{F(q=q^\dagger)-F(q=q_{\min})}{T}]p(E,T)$. The result is:

$$\tau'_{(2)} = \frac{\Delta ET\tau_0}{\sqrt{2\pi}(\gamma T^2 - \Delta E^2)} \exp[\gamma - S_0 + \gamma^2 T^2/(2\Delta E^2) - \Delta E^2/(2T^2)]$$
$$\times \exp[(\gamma T/(\Delta E^2) - 1/T)X] \qquad (5.18)$$

where X is the integration limit for energy E_i, given in the Appendix D.

The thermally averaged escape time from traps is contributed from the combinations of $\tau_{(1)}$, $\tau_{(2)}$, $\tau_{(3)}$, $\tau_{(4)}$, $\tau_{(5)}$, $\tau_{(6)}$, $\tau_{(7)}$, $\tau_{(8)}$ and $\tau_{(9)}$ (see details give in the Appendix D).

To determine the configurational diffusion coefficient, we must discuss also the integration limits on the energy E_i. There are six energy values which control whether the free energy in region of $q < q^\dagger$ and $q > q^\dagger$ is monotonically increasing, decreasing or exhibits a minimum. These six energy values are: $E_1 = -q^\dagger\Delta E^2/T - \alpha T$, $E_2 = -\alpha T$, $E_3 = -\Delta E^2/T - \gamma T$, $E_4 = -q^\dagger\Delta E^2/T - \gamma T$, $E_l = -\sqrt{2S_0\Delta E^2}$ and $E_h = \sqrt{2S_0\Delta E^2}$. The differing orders of these energies values in different temperature ranges naturally define the integration limits for the average escape times.

The algebraic results for escape times in the different temperature ranges are shown in Appendix D.

5.3. ILLUSTRATIVE NUMERICAL RESULTS. — When we increase the size of a RHP, a bilinear form well approximates the configurational entropy, with parameters α and γ (or $q^\dagger$ and $S^\dagger$) adjusted to fit the actual entropy curve as calculated by the methods summarized in Appendix A. To illustrate the results, the average escape time *versus* reduced temperature T/T_G is shown in Figure 3a for 64 mers and Figure 3b for 125 mers (in the case of $\alpha > \gamma$ and $q^\dagger < 0.5$).

We again see that the average escape time still does not follow the Arrhenius law. As temperature is lowered, the escape time exponentially grows leading to effective kinetic trapping. The results obtained by the RSM and ultrametric models become more comparable as chain length increases. In both cases correlation effects result in a significantly smaller escape time than for landscapes without correlations. There exists an N dependent temperature T_A (given in the ultrametric case by Eq. (5.8)) above the thermodynamic glass transition temperature where the dynamics becomes non-activated, and the diffusion time approaches τ_0.

6. Conclusion and Discussion

We have given a detailed analysis of the kinetics of a locally connected generalized random energy model using the random heteropolymer as an illustration. However, the methods used here can be readily applied to study dynamics in any rugged system, *e.g.* spin glasses. We can see that the escape time for a correlated energy landscape is shorter than for an uncorrelated surface, the search time at T_g being reduced but still exponential in the size of the system. The distribution of escape times for an ultrametric RHP was obtained. A second temperature scale emerges in the analyses, analogous to the temperature T_A in Potts spin-glasses above which kinetics is non-activated.

The calculation of the configurational diffusion or escape time on a non-biased correlated energy landscape in this paper can be straightforwardly generalized to study minimally frustrated protein folding, where biasing towards the folded state (minimum frustration) is also taken into account. We leave the detailed treatment of this to issue to future work.

Trap escape and unfolding are mathematically analogous, in that escaping from a large basin or funnel on the energy landscape must involve escaping from a series of smaller basins or funnels, as in the minimally frustrated problem. Escaping from a macro funnel requires escape from a series of micro funnels, which requires escaping from nano funnels, *etc.* until the escape time from the smallest basins or funnels reaches the fundamental microscopic time scale. In this paper, we have renormalized all smaller basin escapes into the microscopic time scale τ_0. A more complete treatment of diffusion on a rugged landscape would require a renormalization group analysis for the diffusion rate in a hierarchy of funnels, which we leave to future work. The result presented here probably overestimates the trapping escape rate while the random energy model result clearly underestimates it. Thus for example the temperature T_A, even for 27-mers, may well exceed T_G^0 as suggested also by replica variational methods [22]. We also point out that topological constraints in longer chains will increase the barriers over the present illustrative results which assume no such prohibitions on kinetics.

The present methods and results, by including correlations, continue beyond the REM in describing diffusion on a rugged landscape. The formalism can flexibly include detailed polymeric effects only sketched here. We believe it should be a valuable next step in understanding complex free energy landscape models both for biopolymers and other systems.

Acknowledgments

Appendix A

We sketch here the calculation of the configurational entropy $S_c(q)$, and refer the interested reader to our earlier work [15].

In the low overlap q limit, that is — for a weakly constrained polymer, the entropy can be decomposed into several terms:

$$S_1 = S_c(\eta) + \Delta S_{\text{contact}} + \Delta S_{\text{mix}} + \Delta S_{\text{AB}}. \tag{A.1}$$

The first term is the entropy of all states given a specified degree of collapse. It is purely a function of packing fraction η [19]

$$S_c(\eta) = N \log \frac{\nu_0}{e} - N \left(\frac{1-\eta}{\eta} \right) \log (1-\eta). \tag{A.2}$$

The second term comes from the loss of entropy involved in formation of specific individual contacts:

$$\Delta S_{\text{contact}} = \frac{3}{2} N q \eta z \left[\log C - 1 + \log(q z \eta) \right] \tag{A.3}$$

where $C = \frac{3}{4\pi} (\frac{\delta\tau}{b^3})^{2/3}$, $\delta\tau$ is the volume corresponding to the interaction range or the bond distance between two residues, b is the persistence length of the polymer.

The third term comes from the combinatorics of choosing which q, z, η, N contacts overlap with a total of $z\eta N$ contacts in the collapsed state. z is the near neighbor coordination number and N is the total number of monomers in the polymer.

$$\Delta S_{\text{mix}} = -Nz\eta \left[q \log q + (1-q) \log(1-q) \right]. \tag{A.4}$$

The fourth term is associated with the fact $Nz\eta - qNz\eta$ contacts of the reference state must not have been formed so that the overlap is no more than q:

$$\Delta S_{AB} = \frac{N}{C} \int_{cq\eta}^{cz\eta} dx \log(1 - x^{3/2}). \tag{A.5}$$

The effect of confinement of the polymer chain is to change the constant in this expression from C to C' where $C' = (12/\pi)(\delta\tau/b^3)^{2/3}$. (This term cancels to some extent the combinatorial term.)

Notice that there is a value of q less than one, q_v, at which $S_1(q_v) = 0$. This value $q_v \sim 1/(z\eta)$ is rather close to where the entropy of a given contact pattern in the mean field would vanish. It is related to the Flory [28] vulcanization value for contacts. Somewhat before this point the polymer is overconstrained and the counting arguments used in deriving the expression for entropy above are invalid. For 27 mer lattices q_v is close to one. But for large polymers, q_v becomes significantly less than one. The weak constraint assumption breaks down near q_v. Important configurations with $q > q_v$ must have significant clustering of the contacts or equivalently *melted* out regions. The dilute contact representation is not good for such large q, so another approximation was developed to take into account the fact that contacts must be formed or broken simultaneously in groups. For $q > q_v$ this melting out is described by changing the representation for low q in terms of an interacting gas of contacts to, at high q, an atomistic description in which the reference state is one where all the contacts are formed. Here entropy arises because contiguous parts of the frozen polymer are melted out and unconstrained. The melted parts carry a certain entropy but there is also a combinatorial entropy associated with the choice of the different places where a given melted piece can occur along the sequence of the polymer. A complete form of the entropy taking into account also the end effects was thus obtained [15].

Fig. 5. — Diagram of the GREM hierarchy used in the ultrametric model.

Appendix B

Consider the GREM hierarchy in Figure 5. Suppose we find a state with energy $E_\imath$ and look for the free energy relative to it as a function of q. We do this in order to find the escape time from i. As in Figure 5,

$$E_\imath = \phi_< + \phi_>^\imath \quad \text{and} \quad E = \phi_< + \phi_>.$$

We wish to find the number of states with energy E and overlap q with state i, $N(E, q, E_i)$. This depends on how much energy came from $\phi_<$, since both $E_\imath$ and E have this contribution. For the ensemble of states with energy $E_\imath$, the distribution of $\phi_<$ is given by the conditional probability distribution

$$
\begin{aligned}
P_q(\phi_< | E_\imath) &= \frac{P(\phi_<, E_\imath)}{P(E_\imath)} \approx \frac{\exp\left(-\dfrac{\phi_<^2}{2\Delta E^2 q}\right) \exp\left(-\dfrac{(E_i - \phi_<)^2}{2\Delta E^2 (1-q)}\right)}{\exp\left(-\dfrac{E_\imath^2}{2\Delta E^2}\right)} \\
&\approx \exp\left(-\frac{(\phi_< - qE_\imath)^2}{2\Delta E^2 q(1-q)}\right).
\end{aligned}
\tag{B.1}
$$

(Notice (B.1) has the correct limits of $\delta(\phi_<)$ and $\delta(\phi_< - E_\imath)$ when $q \to 0$ and $q \to 1$ respectively.) Now suppose we have found a state i with $E_\imath$ and a given contribution $\phi_<$. The log number of states having overlap q with it and energy E are

$$\ln N(E, q, E_\imath) = S_\mathrm{c}(q) - \frac{(E - \phi_<)^2}{2\Delta E^2 (1-q)} \tag{B.2}$$

with $\phi_<$ chosen from the distribution (B.1). Using $1/T = \partial \ln N / \partial E$ gives the free energy

$$F(q, \phi_<) = \phi_< - TS_\mathrm{c}(q) - \frac{\Delta E^2}{2T}(1-q), \tag{B.3}$$

so that the escape time for a given realization of $\phi_<, E_\imath$, is

$$\tau = \tau_0 \exp\left(\frac{F(q^\star) - E_\imath}{T}\right) \tag{B.4}$$

where $F(q^\star)$ is where the free energy has its maximum, given that it is a fluctuating quantity since at a given q, $\phi_<$ fluctuates depending on which state with $E_\imath$ we picked

$$F(q^\star) = \max_q F(q).$$

$q^\star$ also follows a distribution for the different i having E_i, which we will approximate by taking $q^\star$ at its most probable value. (B.1), (B.3) and (B.4) together give the distribution for τ given E_i:

$$P(\tau|E_i)\,d\tau \approx \frac{d\tau}{\tau} \exp - \frac{\left(T\ln\frac{\tau}{\tau'_\bullet} + (1-q^\star)\left(E_i + \frac{\Delta E^2}{T}\right)\right)^2}{2\Delta E^2 q^\star(1-q^\star)} \tag{B.5}$$

where $\tau' = \tau_0 \exp - S_c(q^\star)$. The distribution of escape times is obtained by averaging (B.5) over the thermal distribution of energies

$$P_T(\tau) = \int_{E_{GS}^-}^{E_{GS}^+} dE_i\, P_T(E_i)\, P(\tau|E_i) \tag{B.6}$$

where $E_{GS}^\pm = \pm\sqrt{2S_0\Delta E^2}$ and

$$P_T(E_i) \approx \exp - \frac{\left(E_i + \frac{\Delta E^2}{T}\right)^2}{2\Delta E^2}$$

The integrand is extensive and steepest descents can be taken, with the largest contribution to $P(\tau)$ coming from

$$E_i^\star = -\frac{\Delta E^2}{T} - T\ln\frac{\tau}{\tau'} \tag{B.7}$$

(the dominant energies contributing to the escape time are always lower than the thermal energy) and the distribution of escape times at temperature T is then given by

$$P_T(\tau) \approx \frac{1}{\tau} \exp - \frac{\left(T\ln\frac{\tau}{\tau'}\right)^2}{2\Delta E^2(1-q^\star)}. \tag{B.8}$$

Note that as $q^\star \to 1$, i.e. diffusion becomes a non-activated, downhill process, $E_i^\star \to -\Delta E^2/T$ and $P_T(\tau) \to \delta(\tau - \tau_0)$.

Equations (B.1, B.3) together allow one to consider the probability distribution of the free energy as a function of q

$$P_q(F) \approx \exp - \frac{1}{2\Delta E^2 q(1-q)}\left(F - qE_i + TS_c(q) + \frac{\Delta E^2}{2T}(1-q)\right)^2, \tag{B.9}$$

with the mean and most probable free energy function $F^\star(q)$ given by equation (2.11).

Appendix C

The high temperature free energy (4.1) for $T > \sqrt{2}T_G^0$ is

$$F_m = \sqrt{\frac{S_0\Delta E^2}{2}}\left(-x - \frac{1}{x} + q(x-2y) + \frac{q^2}{x}\right) \tag{C.1}$$

with a minimum at

$$q^\ddagger = \frac{x}{2}(2y-x)\cdot \tag{C.2}$$

The upper bound to energies with activated barriers occurs when $q^\ddagger = 0$ or

$$y = \frac{x}{2}$$

Using the thermal distribution of energies

$$P(y, T) = \frac{1}{\sqrt{2\pi \Delta E^2}} \exp{-S_0 \left(y - \frac{1}{x}\right)^2}$$

in equation (4.2) gives for the mean escape time

$$\langle \tau(x) \rangle = \tau_0 \sqrt{\frac{S_0}{\pi}} \exp{S_0 \left(\frac{x^2}{4} - \frac{1}{x^2}\right)} \times \int_{x/2}^{1} dy \, \exp{-S_0 \left(x - \frac{2}{x}\right) y} \qquad \text{(C.3)}$$

which for $x > \sqrt{2}$, is of $\mathcal{O}\left(\mathrm{e}^{-N}\right)$. So (for a large system) there are no macroscopic barriers above $\sqrt{2}T_G^0$ (even though states with barriers exist, at these high temperatures they are not macroscopically populated).

Note that at $T = \sqrt{2}T_G^0$, the mean escape time scales as a power law with system size

$$\left\langle \tau\left(\sqrt{2}T_G^0\right)\right\rangle = \tau_0 \sqrt{\frac{S_0}{\pi}} \left(1 - \frac{1}{\sqrt{2}}\right) \sim \mathcal{O}\left(N^{1/2}\right). \qquad \text{(C.4)}$$

As long as the escape barrier is calculated from the melted free energy F_m in (4.1), expression (C.3) is true. This condition is that the free energy minimum is always in the melted region of the piece-wise free energy of (4.1), or that $q^\ddagger (y = 1) < q_G(x)$, or $x > x^\# = \left(1 + \sqrt{7}\right)/3$. For $x < \sqrt{2}$ however, expression (C.3) dictates the escape times scale exponentially with system size $\langle \tau(T) \rangle \sim \mathcal{O}\left(\mathrm{e}^N\right)$.

For low temperatures such that $1 < x < x^\#$, there is an energy $y^\ddagger(x) E_{GS}^0$ below which the free energy minimum is evaluated from the frozen formula F_f in (4.1). $y^\ddagger(x)$ is given by the condition that the minima of F_m and F_f occur at the same place, *i.e.* $q_m^\ddagger = q_f^\ddagger$, or

$$y^\ddagger(x) = \frac{3x^2 - 2}{2x} \qquad \text{(C.5)}$$

This value of energy splits the integral in (4.2) into two parts, thus

$$\frac{\langle \tau(x) \rangle}{\tau_0} = H + L \qquad \text{(C.6)}$$

where H is the integral over higher energies $x/2 < y < y^\ddagger(x)$, when the barrier is in the unfrozen regime:

$$H = \frac{1}{\sqrt{\pi S_0}} \left(\frac{2}{x} - x\right)^{-1} \exp{S_0 \left(\frac{x^2}{4} - \frac{1}{x^2}\right)} \times \left[\exp{S_0 \left(\frac{2}{x} - x\right) y}\right]_{y = x/2}^{y = \frac{3x^2 - 2}{2x}} \qquad \text{(C.7)}$$

Note H becomes smaller at lower temperatures, and $H \to 0$ when $x \to 1$.

L is the contribution to the escape time from low energies $y^\ddagger(x) < y < 1$, and the barrier must be calculated from the minimum of the frozen free energy (F_f in (4.1)), to the maximum of the melted free energy F_m at $q = 0$. This calculation gives

$$L = \sqrt{\frac{S_0}{\pi}} \int_{y = \frac{3x^2 - 2}{2x}}^{1} dy \, \exp{S_0 \left[\frac{2}{x} g(y) - 1 - \frac{1}{x^2} - \left(y - \frac{1}{x}\right)^2\right]} \qquad \text{(C.8)}$$

where

$$g(y) = \frac{2\sqrt{2}}{27} \left(6 - y^2 - y\sqrt{6 + y^2}\right)\left(3 + y^2 + y\sqrt{6 + y^2}\right)^{1/2}$$
$$+ \frac{1}{18}y\left(4y^2 - 6 + 4y\sqrt{6 + y^2}\right).$$

For the range of temperatures $1 < x < x^{\#}$, the largest contribution to $\langle \tau(T) \rangle$ in the integral is from $y = 1$, or the ground state energy E_{GS}^0. Approximating the integral by steepest descents gives

$$L \cong \frac{1}{2\sqrt{\pi S_0}}\left(\frac{c}{x} - 1\right)^{-1}\exp 2S_0\left(\frac{g(1) - c + 1}{x} - \frac{1}{x^2}\right) \times \left[\exp 2S_0\left(\frac{c}{x} - 1\right)y\right]_{y = \frac{3x^2 - 2}{2x}}^{y = 1} \quad \text{(C.9)}$$

where

$$g(1) = \frac{1}{9}\left(2\sqrt{7} - 1\right) + \left(1 - \frac{1}{9}\left(2\sqrt{7} - 1\right)\right)\left(1 + \frac{1}{9}\left(2\sqrt{7} - 1\right)\right)^{1/2} \cong 1.113$$

$$c = \frac{2}{9}\left(6 + \frac{15}{\sqrt{7}} - \sqrt{\frac{2}{7}}\left(4 + \sqrt{7}\right)^{1/2}\left(1 + \sqrt{7}\right)\right) \cong 1.477.$$

L and $\langle \tau(T) \rangle$ continue to increase as the temperature is lowered, until at T_G^0 ($x = 1$), $H = 0$ and

$$\langle \tau(T_G^0) \rangle \cong \frac{\tau_0}{2\sqrt{\pi S_0}}\left(\frac{1}{c - 1}\right)\exp 2S_0\left(g(1) - 1\right)$$
$$\approx \frac{\tau_0}{\sqrt{\pi S_0}}\exp\left(0.225 \times S_0\right). \quad \text{(C.10)}$$

Note that the coefficient in the exponent as well as the prefactor reduce the escape time from the REM value of $\tau_0 \exp S_0$.

Appendix D

For the piece-wise linear entropy function, the minimum of the free energy can be located in different regions of q depending on the chosen E_i and temperature T. In region 1, if $q_{min}^1 < 0$, this corresponds to $E_i > E_2 = -T\alpha$, the free energy function is monotonically increasing between 0 and $q^{\dagger}$. On the other hand, if $0 < q_{min}^1 < q^{\dagger}$, this corresponds with $E_1 = -T\alpha - q^{\dagger}\Delta E^2/T < E_i < -T\alpha$, the free energy now has a minimum between 0 and $q^{\dagger}$. If $q_{min}^1 > 1$, this corresponds with $E_i < E_1$, the free energy now is monotonically decreasing between 0 and $q^{\dagger}$. In regime 2, if $q_{min}^2 < q^{\dagger}$, this corresponds with $E_i > E_4 = -T\gamma - q^{\dagger}\Delta E^2/T$, the free energy now monotonically increases. If $q^{\dagger} < q_{min}^2 < 1$, this corresponds with $E_3 = -T\gamma - \Delta E^2/T < E_i < E_4$, the free energy has a minimum. If $q_{min}^2 > 1$, this corresponds to $E_i < E_3$, the free energy is monotonically decreasing. So there are 3 possible situations for the free energy in the low q (regime 1) with monotonically decreasing, minimum, monotonically increasing behaviour and 3 possible situations for free energy at high q (regime 2) with monotonically decreasing, minimum, monotonically increasing behaviour.

There are altogether 9 possible combinations of situations for the free energy as a whole in high and low q regime. We discuss these below.

The 9 situations for the free energy are shown in Figure 4. We term them situations (1), (2), (3), (4), (5), (6), (7), (8), (9) respectively. The only differences among these 9 different
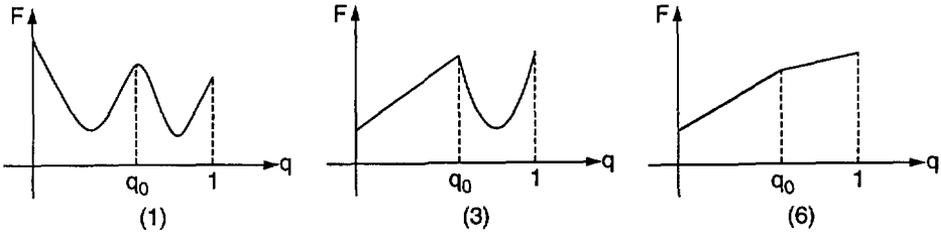
Fig. 6. — The free energy profiles of different energy ranges *versus* overlap or fraction of contacts q are shown for 64 mers with the piece-wise linear approximation for the entropy of temperature range $T_g < T < T_0$.

shapes of the free energies are the ranges of the chosen energy E_i. The ranges of E_i for the free energy of situation (1) are $E_1 < E_i < E_2$ and $E_3 < E_i < E_4$; the ranges of E_i for the free energy of situation (2) are $E_i < E_1$ and $E_3 < E_i < E_4$; the ranges of E_i for the free energy of situation (3) are $E_i > E_2$ and $E_3 < E_i < E_4$; the ranges of E_i for the free energy of situation (4) are $E_1 < E_i < E_2$ and $E_i > E_4$; the ranges of E_i for the free energy of situation (5) are $E_i < E_1$ and $E_i > E_4$; the ranges of E_i for the free energy of situation (6) are $E_i > E_2$ and $E_i > E_4$; the ranges of E_i for the free energy of situation (7) are $E_1 < E_i < E_2$ and $E_i < E_3$; the ranges of E_i for the free energy of situation (8) are $E_i < E_1$ and $E_i < E_3$; the ranges of E_i for the free energy of situation (9) are $E_i > E_2$ and $E_i < E_3$.

The four energy scales which control whether the free energy in region of $q < q^\dagger$ and $q > q^\dagger$ is monotonically increasing, decreasing or has a minimum are: $E_1 = -q^\dagger \Delta E^2 / T - \alpha T$, $E_2 = -\alpha T$, $E_3 = -\Delta E^2 / T - \gamma T$, $E_4 = -q^\dagger \Delta E^2 / T - \gamma T$. As discussed earlier at a given Q, the heteropolymer also has a low and a high energy cut off given by: $E_l = \overline{E(Q)} - \sqrt{2S(Q)\Delta E^2}$ and $E_h = \overline{E(Q)} + \sqrt{2S(Q)\Delta E^2}$. We will simplify the algebra by assuming from now on that $\overline{E(Q)} = 0$, $\Delta E = \Delta E$. This simplification changes nothing significant since we only care about general diffusion or escaping time of the locally connected correlated energy landscape here. These energy ranges depend on temperature. By comparing E_1, E_2, E_3 and E_4, we see that different orders of these energy values correspond with different temperature ranges.

For the purpose of this paper, we concentrate our discussion on the diffusion time with the parameters appropriate for the 64 unit and 125 unit polymers where $q^\dagger < 0.5$ and $\alpha > \gamma$ (the slope of the initial entropy loss for low q is greater than the slope of the final entropy loss for high q with the parameters of the corresponding random heteropolymer). We found that for 64 mers $E_l > E_3$ always; $E_l > E_1$ always; $E_l < E_2$ at temperature T_0 where T_0 is determined by the equation $E_l = E_2$. In other words, we see that in the temperature range $T_g < T < T_0$, $E_l < E_2 < E_4 < E_h$; in the temperature range $T_0 < T$, $E_l < E_4 < E_h$. For 125 mers, we found that for $T > T_g$, $E_l > E_3$ always; $E_l > E_1$ always; $E_l > E_2$ always; then we obtain that $E_l < E_4 < E_h$.

As mentioned earlier, for each temperature zone, there are 6 energy values (E_l, E_1, E_2, E_3, E_4, E_h) which define different regions of energies each giving a different shape of the free energies. This provides naturally the integration limit for the averaged escape times. These different shapes of the free energies in various of temperature ranges are shown in Figures 6, 7 and 8.

For 64 mers, $q^\dagger < 0.5$ and $\alpha > \gamma$, we obtain the thermally averaged escape time for temperature $T_g < T < T_0$, $\langle \tau^1 \rangle = \tau_1^1 + \tau_2^1 + \tau_3^1$ where the upper case of τ refers to the temperature range and the lower case of τ refers to sequentially the energy regions (low to high) which
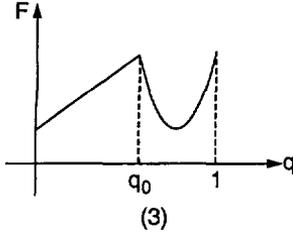
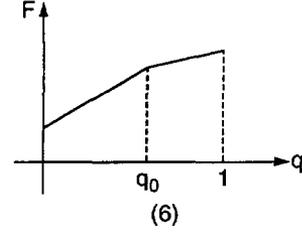Fig. 7.   Fig. 8.

Fig. 7. — The free energy profiles of different energy ranges *versus* overlap or fraction of contacts q are shown for 64 mers with the piece-wise linear approximation for the entropy of temperature range $T_0 < T$.

Fig. 8. — The free energy profiles of different energy ranges *versus* overlap or fraction of contacts q are shown for 125 mers with the piece-wise linear approximation for the entropy of temperature range $T_g < T$.

give the integration limits for the escape time. $\tau_{(1)}$, $\tau_{(2)}$, $\tau_{(3)}$, $\tau_{(4)}$, $\tau_{(5)}$, $\tau_{(6)}$, $\tau_{(7)}$, $\tau_{(8)}$ and $\tau_{(9)}$ correspond to expressions for $\tau'_{(1)}$, $\tau'_{(2)}$, $\tau'_{(3)}$, $\tau'_{(4)}$, $\tau'_{(5)}$, $\tau'_{(6)}$, $\tau'_{(7)}$, $\tau'_{(8)}$ and $\tau'_{(9)}$ mentioned in Section 4.1 with the integration limits for different temperature and energy ranges now explicitly put in:

$$
\begin{aligned}
\tau_1^1 &= \tau_{(1)} \\
&= \frac{\Delta E T \tau_0}{\sqrt{2\pi}(\gamma T^2 - \Delta E^2(1 - q^\dagger))} \exp[\gamma^2 T^2/(2\Delta E^2) + \gamma q^\dagger + q^{\dagger^2}\Delta E^2/(2T^2)] \\
&\quad \times \exp[-\Delta E^2/(2T^2)](\exp[(\gamma T/\Delta E^2 + (q^\dagger/T - 1/T))(-\alpha T)] \\
&\quad - \exp[(\gamma T/\Delta E^2 + (q^\dagger/T - 1/T))(-\sqrt{2S_0\Delta E^2})])
\end{aligned}
\tag{D.1}
$$

$$
\begin{aligned}
\tau_2^1 &= \tau_{(3)} \\
&= \frac{\Delta E T \tau_0}{\sqrt{2\pi}(\gamma T^2 - \Delta E^2(1 - q^\dagger))} \exp[\gamma^2 T^2/(2\Delta E^2) + \gamma q^\dagger + q^{\dagger^2}\Delta E^2/(2T^2)] \\
&\quad \times \exp[-\Delta E^2/(2T^2)](\exp[(\gamma T/\Delta E^2 + (q^\dagger/T - 1/T))(-q^\dagger\Delta E^2/T - \gamma T)] \\
&\quad - \exp[(\gamma T/\Delta E^2 + (q^\dagger/T - 1/T))(-\alpha T)])
\end{aligned}
\tag{D.2}
$$

$$
\begin{aligned}
\tau_3^1 &= \tau_{(6)} \\
&= \frac{1}{2}\tau_0(\mathrm{erf}[\frac{\Delta E^2 + T\sqrt{2S_0\Delta E^2}}{\sqrt{2}\Delta E T}] - \mathrm{erf}[\frac{\Delta E^2 - \gamma T^2 - q^\dagger\Delta E^2}{\sqrt{2}\Delta E T}]).
\end{aligned}
\tag{D.3}
$$

For 64 mers, in the temperature range $T_0 < T$, the thermally averaged escape time $\langle \tau^2 \rangle = \tau_1^2 + \tau_2^2$:

$$
\begin{aligned}
\tau_1^2 &= \tau_{(3)} \\
&= \frac{\Delta E T \tau_0}{\sqrt{2\pi}(\gamma T^2 - \Delta E^2(1 - q^\dagger))} \exp[\gamma^2 T^2/(2\Delta E^2) + \gamma q^\dagger + q^{\dagger^2}\Delta E^2/(2T^2)] \\
&\quad \times \exp[-\Delta E^2/(2T^2)](\exp[(\gamma T/\Delta E^2 + (q^\dagger/T - 1/T))(-q^\dagger\Delta E^2/T - \gamma T)] \\
&\quad - \exp[(\gamma T/\Delta E^2 + (q^\dagger/T - 1/T))(-\sqrt{2S_0\Delta E^2})])
\end{aligned}
\tag{D.4}
$$

$$\tau_2^2 = \tau_{(6)}$$
$$= \frac{1}{2}\tau_0(\mathrm{erf}[\frac{\Delta E^2 + T\sqrt{2S_0\Delta E^2}}{\sqrt{2}\Delta ET}] - \mathrm{erf}[\frac{\Delta E^2 - \gamma T^2 - q^\dagger\Delta E^2}{\sqrt{2}\Delta ET}]). \qquad (D.5)$$

For 125 mers, in the temperature range $T > T_{\mathrm{g}}$, the thermally averaged escape time $\langle\tau\rangle = \tau_1 + \tau_2$:

$$\tau_1 = \tau_{(3)}$$
$$= \frac{\Delta ET\tau_0}{\sqrt{2\pi}(\gamma T^2 - \Delta E^2(1 - q^\dagger))}\exp[\gamma^2 T^2/(2\Delta E^2) + \gamma q^\dagger + q^{\dagger^2}\Delta E^2/(2T^2)]$$
$$\times \exp[-\Delta E^2/(2T^2)](\exp[(\gamma T/\Delta E^2 + (q^\dagger/T - 1/T))(-q^\dagger\Delta E^2/T - \gamma T)]$$
$$- \exp[(\gamma T/\Delta E^2 + (q^\dagger/T - 1/T))(-\sqrt{2S_0\Delta E^2})]) \qquad (D.6)$$

$$\tau_2 = \tau_{(6)}$$
$$= \frac{1}{2}\tau_0(\mathrm{erf}[\frac{\Delta E^2 + T\sqrt{2S_0\Delta E^2}}{\sqrt{2}\Delta ET}] - \mathrm{erf}[\frac{\Delta E^2 - \gamma T^2 - q^\dagger\Delta E^2}{\sqrt{2}\Delta ET}]). \qquad (D.7)$$

References

[1] Frauenfelder H., Sligar S. and Wolynes P.G., *Science* **254** (1991) 1598 and references therein; Frauenfelder H., Wolynes P.G., *Phys. Today* **47** (1994) 58; Wolynes P.G., in "Spin Glass and Biology", Stein, Ed., (World Sci entific Pub., 1992) pp. 225-259.

[2] Mézard M., Parisi G., Virasoro M.A., Spin Glass Theory and Beyond (World Scientific Pub., 1987).

[3] Bryngelson J.D. and Wolynes P.G., *Proc. Natl. Acad. Sci USA* **84** (1987) 7524; *J. Phys. Chem.* **93** (1989) 6902.

[4] Onuchic J.N., Wolynes P.G., Luthey-Schulten Z.A. and Socci N.D., *Proc. Natl. Acad. Sci.* **92** (1995) 3626.

[5] Socci N.D., Onuchic J.N. and Wolynes P.G., *J. Chem. Phys.* **104** (1996) 5860-5868.

[6] Goldstein R.A., Luthey-Schulten Z.A. and Wolynes P.G., *Proc. Natl. Acad. Sci. USA* **89** (1992) 9029.

[7] Bouchaud J.P. and Dean D.S., *J. Phys. I France* **5** (1995) 265-286.

[8] DeDominicis C., Orland H., Lainée F., *J. Phys. Lett. France* **46** (1985) L463; Koper G.J.M., Hilhorst H.J., *Europhys. Lett.* **3** (1987) 1213.

[9] Bryngelson J.D., Onuchic J.N., Socci N.D., Wolynes P.G., *Proteins Struct. Funct. Genet.* **21** (1995) 167 and references therein.

[10] Wolynes P.G., Onuchic J.N., D. Thirumalai D., *Science* **267** (1995) 1619; Chan H.S., Dill K.A., *J. Chem. Phys.* **100** (1994) 9238; Abkevich V.I., Gutin A.M., Shakhnovich E.I., *J. Chem. Phys.* **101** (1994) 6052.

[11] Derrida B., *Phys. Rev. B* **24** (1981) 2613.

[12] Gross D.J., Kanter I. and Sompolinsky H., *Phys. Rev. Lett.* **55** (1985) 304.

[13] Shakhnovich E.I. and Gutin A.M., *Europhys. Lett.* **8** (1989) pp. 327-332.

[14] Sasai M. and Wolynes P.G., *Phys. Rev. Lett.* **65** (1990) 2740; Shakhnovich E.I. and Gutin A.M., *Biophys. Chem.* **34** (1989) 187; Garel T. and Orland H., *Europhys. Lett.* **6** (1988) 307.

[15] Plotkin S., Wang J., Wolynes P.G., *Phys. Rev. E* **53** (1996) 6271.

[16] Plotkin S., Wang J., Wolynes P.G., *J. Chem. Phys* **106** (1997) 2932-2948.

[17] Levinthal C., Proceedings of Meeting at Allerton House, P. DeBrunner, J. Tsibris and E. Munck, Eds. (Univ. of Illinois Press, 1969) p. 22.

[18] Derrida B., *J. Phys. Lett.* **46** (1985) L401; Derrida B. and Gardner E., *J. Phys. C* **19** (1986) 2253.

[19] Bryngelson J.D. and Wolynes P.G., *Biopolymers* **30** (1990) 177.

[20] Kirkpatrick T.R. and Wolynes P.G. *Phys. Rev. B* **36** 8552(1987) 8552; Kirkpatrick T.R., Thirumalai D. and Wolynes P.G., *Phys. Rev. A* **40** (1989) 1045.

[21] Kirkpatrick T.R. and Wolynes P.G., *Phys. Rev. A* **35** (1987) 3072.

[22] Takada S. and Wolynes P.G., *Phys. Rev. E* (to appear).

[23] Saven J.G., Wang J. and Wolynes P.G., *J. Chem Phys* **101** (1994) 11037-11043; Wang J., Saven J.G. and Wolynes P.G., *J. Chem. Phys.* **105** (1996) 11276-11284.

[24] Hagen S.J., Hofrichter J., Szabo A. and Eaton W.A., *Proc Natl. Acad. Sci.* **93** (1996) 11615-11617.

[25] Unless otherwise indicated, all entropies have dimensions of Boltzmann's constant k_B.

[26] Unless otherwise indicated, all temperatures have units of energy, with Boltzmann's constant k_B included in the definition of T.

[27] Koretke K.K., Luthey-Schulten Z. and Wolynes P.G., *Protein Sci.* **5** (1996) 1043-1059.

[28] Flory P.J, *Chem. Rev.* **35** (1944) 51; *J. Chem. Phys.* **18** (1950) 108.