Chapter 5
On Relaxation to the Global Minimum

5.1 Introduction

How difficult is it to go from a random point on a PES to the global minimum? The answer, of course, depends upon the system. Many proteins are able to fold reversibly to the physiologically active native structure from unfolded states. However, a random amino acid sequence is very unlikely to be able to fold to a unique structure. Similarly, in condensed matter physics there are substances for which it is hard to prevent crystallization as the liquid is cooled, and others which instead are likely to form glasses.

The seminal analysis of the global optimization problem in the context of protein folding was made by Levinthal[226]. This author estimated the number of configurations a typical protein could adopt (3^N where N is the number of amino acids) and noted that even if protein configurations could be sampled at a rate of, say, 10^{13} per second, it would take longer than the age of the universe to find the native structure if this sampling was simply random. The result of this simple calculation markedly contradicts the actual properties of proteins, which can generally find the native structure from an unfolded state on time scales of the order of seconds or less. This well-known contradiction has come to be known as Levinthal's paradox. A similar paradox can be formulated for a cluster. For example, the number of geometrically distinct minima on the PES of a 55-atom cluster interacting via the Lennard-Jones potential has been estimated to be of the order of 10^{21}, yet the cluster can rapidly find the Mackay icosahedral global minimum from the liquid-like state[110,111]. The existence of 'magic number' peaks in the mass spectra of rare gas clusters[47,48] indicates that these systems can also locate certain global minima very efficiently.

A more rigorous approach to defining the effort involved in finding the global minimum can be obtained using computational complexity theory[227], which categorizes problems into classes of similar difficulty. The most relevant class is that of NP-hard problems, for which there is no known algorithm that is guaranteed to find the solution within polynomial time, effectively rendering NP-hard problems intractable for large sizes. It has been shown that finding the global minimum of a protein[228,229] or a cluster[93] is NP-hard. Again, there seems to be a paradox between the NP-hardness of protein folding and the actual behaviour of proteins. However, it should be noted that computational complexity theory provides a worst case analysis. First, the proofs apply to general categories of problems and not necessarily to every instance of a problem. Although there may be no efficient method for finding the global minimum of a general heteropolymer, biological proteins may represent a subset whose properties have evolved so that they can fold. Secondly, the criterion for a successful algorithm is very stringent; one must not only find the global minimum, but prove that this it is actually the global minimum. In contrast, proteins do not necessarily have to fold to the global potential energy minimum, but instead must have a high probability of folding to the native structure which probably corresponds to one of the lowest energy minima on the PES. Similarly, the heuristic global optimization methods, such as simulated annealing[107] and genetic algorithms[230,231] are not rigorously guaranteed to find the global minimum. For example, these methods can find the Mackay icosahedron for the 55-atom Lennard-Jones cluster, which is widely accepted to be the global minimum even though this has never been rigorously proved.

Some have tried to resolve Levinthal's paradox by hypothesizing a reduction in the number of states that the protein has to search through. Typically it is argued that the protein first collapses rapidly to a compact conformation and that the subsequent search is through the greatly reduced space of these compact states[232–234]. Although in some cases this might so reduce the number of states that Levinthal's paradox no longer applies, it cannot in general provide a complete answer. For instance, the number of minima corresponding to the liquid-like state of a 55-atom Lennard-Jones cluster is of the order of 10^{12} (§3.6). This is much less than the total number of minima on the PES but the ease with which the global minimum can be found is still incompatible with a random search through this reduced configuration space. In fact, when the cluster has an energy or temperature in the melting region, it can switch back and forth between the Mackay icosahedral global minimum and the liquid-like states on a time scale of the order of nanoseconds (using parameters appropriate for argon)[110,111].

Seeking to resolve Levinthal's paradox by reducing the search space may be unproductive in a more fundamental sense, for it ignores a basic fallacy in Levinthal's paradox: namely the assumption that each point in configuration space is equally likely. Levinthal was effectively assuming that the PES is totally flat. However, the topography of a protein PES is far from flat, involving many 'mountain ranges' and 'valleys', thus making some configurations more likely than others. (In the canonical ensemble the low potential energy configurations have larger Boltzmann weights, and in the microcanonical ensemble they have a larger momentum density of states.) To illustrate this point Zwanzig *et al.* produced a simple model of a protein with an energetic bias towards the native structure which could find the global minimum on physically reasonable time scales[235]. However, this particular model problem belongs[229] to the class of problems P, which are tractable in polynomial time, not to the class NP.

In lattice model studies of proteins, sequences that fold are often compared with those that do not. Some studies have suggested that one of the distinctive characteristics of the folding sequences is a significant energy gap between the global minimum and the next lowest energy structure[233,234,236] and this criterion has been used to design sequences that fold rapidly[237,238]. One particularly comprehensive study of a 27-unit chain led Sali *et al.* to conclude that the necessary and sufficient condition for a folding sequence is that the native state be a pronounced global minimum[233,234]. However, this conclusion has been criticized both because the correlation between the energy gap and folding ability is rather weak and because the criterion concentrates on only a very small part of the PES, whereas folding ability surely depends on the entire energy landscape[239]. Interestingly, Sali *et al.*'s results actually show a stronger correlation between the folding ability and the temperature, T_x, at which the global minimum has an equilibrium probability of 0.8, if only compact states are considered.

There are further problems when trying to apply the energy gap criterion to real proteins, since for a real protein there are many minima on the PES which involve only small perturbations to the native structure and have very similar potential energies[240]. In contrast, for the lattice model considered by Sali *et al.* any change to the native structure involves a much larger perturbation and can give rise to a large increase in the energy. Therefore, when applied to more complex models and real proteins, it is probably best to interpret the energy gap criterion as the energy gap between the native state and the lowest energy structurally distinct non-native state. The higher the energy of the non-native states relative to the global minimum the less likely they are to act as kinetic traps. Support for this reinterpretation has been found from an annealing study of simple off-lattice proteins[241].

Other studies of lattice models have come to very different conclusions from Sali *et al.* In particular, Klimov and Thirumalai report no correlation between the folding time and the energy gap between the two lowest energy states[242]. Instead, they found that the folding time decreased as the ratio of the folding temperature, T_f, (at which the native structure becomes the thermodynamically most stable state) to the collapse transition temperature, T_θ, (at which the protein transforms from an extended to a compact state) increased. The source of the differences between this study and that of Sali *et al.* is not clear since they involved very similar systems.

An alternative approach to understanding protein folding has been put forward by Wolynes and coworkers[243–247]. It is based upon a statistical characterization of the free energy landscape, and has its origin in spin-glass theory. This approach was combined with the idea of a 'folding funnel', a set of convergent pathways which lead to the global minimum, making it kinetically accessible from the ensemble of misfolded states[248]. A protein that can fold would have a single folding funnel, whereas a random heteropolymer is likely to have numerous funnels leading to structurally different low energy states. The theory describes the global properties of the free energy surface in terms of a few parameters: the ruggedness of the surface, the extent of the funnel (the size of the configurational space), and the potential energy gradient towards the native state. A rugged PES has many minima with large barriers between them. For a protein to be able to fold the funnel would need a sufficiently large potential energy gradient to direct the folding protein to the native structure.

In the above model, the slope of the funnel can also be related to the energy gap between the native state and the disordered collapsed structures, (a different energy gap to that used by Sali *et al.*). By maximizing this energy gap the ratio of T_f (defined above) to the glass transition temperature, T_g, (at which the dynamics dramatically slow down because of kinetic barriers to escape local minima on the PES) is maximized. A large ratio T_f/T_g ensures that the native state is the most stable state at temperatures where the dynamics are still fast, allowing the protein to reach the native state rather than becoming trapped in a local minimum. This criterion can also explain the performance of global optimization methods, such as simulated annealing, which attempt to follow the global free energy minimum as a system is cooled. Such methods are likely to fail when the global minimum of the free energy changes at a temperature below T_g. This approach can also help us to understand some aspects of the results for lattice model proteins that we noted earlier. The T_x that was used by Sali *et al.* has a very similar definition to T_f; both are measures of the stability of the native state of the protein. It is, therefore, unsurprising that increasing T_x (and hence T_f/T_g, if T_g is fairly independent of sequence) increases the ability of the protein to reach the native state. Also, Thirumalai has emphasized[249] the relationship between the values of the ratios T_f/T_θ and T_f/T_g.

Another equivalent formulation of the above ideas is in the 'principle of minimal frustration'. Proteins that have a single folding funnel are likely to have a native state in which all the interactions between residues are favourable. In contrast, in compact configurations of a random heteropolymer there are likely to be residues brought together with conflicting interactions—a random heteropolymer is 'frustrated'. Minimizing the frustration leads to a low energy native structure and consequently to a large value for T_f. These ideas not only provide an elegant theoretical understanding of protein folding, but by using experimental information to estimate the characteristic parameters of the energy landscape for real proteins, they allow a connection to be made between real proteins and lattice models[250].

These theories developed in the context of protein folding can also be useful in understanding the dynamics of systems in the realm of condensed matter physics. As we noted in §4.4, simple liquids have been described as frustrated systems: they have significant polytetrahedral character, yet all space cannot be packed with regular tetrahedra. Therefore, close-packed structures, which are packings of both tetrahedra and octahedra, are the densest and lowest energy structures. Frank was the first to suggest that this structural difference between the solid and the liquid was responsible for the large degree of supercooling that could be achieved for liquid metals[188]. The widest funnels on the PES in such systems lead to glassy minima rather than to close-packed structures.

The packing constraints, and hence the degree of frustration, can be changed by introducing positive curvature into space. In fact for a space of appropriate positive curvature, the lowest energy, highest density structure is a perfect polytetrahedral packing[191] called polytope $\{3, 3, 5\}$. Straley performed a series of simulations which showed that crystallization occurs much more rapidly in this curved space than in Euclidean space, elegantly demonstrating the effect of frustration on the dynamics[200]. For the curved space there is a single funnel on the PES which leads down through the liquid states to the global minimum.

Similarly, for small Lennard-Jones clusters global optimization strategies, such as genetic algorithms, are able to find the global minimum based on icosahedra from the mass of liquid-like structures fairly easily[73] because icosahedral structures have significant polytetrahedral character. However, there are a number of sizes at which small Lennard-Jones clusters do not have an icosahedral global minimum. For example, the global minimum of a 38-atom cluster is the face-centred-cubic truncated octahedron[70,71] and the lowest energy structures for clusters with 75, 76, 77, 102, 103 and 104 atoms are probably decahedral[71,72]. These cases are much harder for a global optimization algorithm to find since the structures have a much greater close-packed character. In fact the decahedral clusters have never been found yet by an unbiased global optimization method*. It is much harder for the cluster to find the funnel associated with the decahedral minima, compared to the wide funnel that leads down to the icosahedral minima.

Interestingly, in a study of small Lennard-Jones clusters (19 atoms and less) which used global optimization methods based upon the simulated annealing of the classical density distribution, the efficiency was found to be correlated with the energy gap between the global minimum and the next lowest energy state[251]. However, it may be dangerous to draw general conclusions from this result. First, the difference in dynamic behaviour between proteins that can and cannot fold is much greater than that between clusters of different sizes; all the clusters considered in the above study can reach the global minimum relatively easy. Secondly, for these very small systems, much of the thermodynamics is determined by the low energy minima on the PES and the energy gap between the two lowest energy structures is likely to correlate with the melting temperature. For example, the 'melting' transition of the 13-atom cluster is associated with transitions between the icosahedron and the set of the defective icosahedral minima which are next lowest in energy[128]. However, for larger clusters melting is associated with a transition between the low energy minima and the numerous high energy amorphous minima, and the melting temperature is now likely to be correlated with the energy gap between the global minimum and this band of 'liquid-like' minima.

The complexity of proteins means that, inevitably, one has to be satisfied with

*The 75-atom Marks decahedron, though, has been found by a global optimization method developed by David Wales where a decahedral seed is introduced at the centre of the cluster.

a coarse-grained picture of the dynamics involved in protein folding. Studies of small clusters, therefore, could offer particularly important insights since their size allows a much more detailed analysis of the PES[252] and the time scale of the transition from the disordered liquid-like states to a unique solid-like structure is short enough to be probed by computer simulation. Initial studies exploring the relationship between the dynamic behaviour of clusters and the PES have already produced interesting results[252–254]. For example, a comparison of potassium chloride and argon clusters has shown that the greater ability of potassium chloride clusters to crystallize into solid-like structures from the liquid is due to the lower energy barriers and the greater potential energy gradient along series of reaction paths which lead to rock-salt type structures[252]. A more quantitative study sought to understand the behaviour of a 19-atom Lennard-Jones cluster based upon large samples of minima and transition states. The dynamical behaviour was found by calculating transition rates between minima, and then constructing and solving a master equation to describe the flow of probability through configuration space[253,254]. The latter results, however, did not include corrections for the fact that the samples represent only a small fraction of the total number of stationary points on the PES.

In this chapter we follow a similar approach, proceeding from a complete description of a model PES in terms of its minima and transition states to the thermodynamic and dynamic behaviour. We concentrate on the ability of the system to find the global minimum and how this is related to the thermodynamics through Landau entropy profiles. This approach allows us to systematically investigate the effects of various topographic features. In §5.2 we describe the basic PES and the methods for calculating the thermodynamic and dynamic behaviour. In §5.3 we present the results for the PES with a single folding funnel, and in §5.4 we consider the results for the modified surfaces. Finally in §5.5, we summarize the main conclusions from this work. Throughout the chapter we attempt to show how our approach can illuminate and provide a framework for understanding the results already known for specific systems, such as proteins, clusters and glasses.

5.2 Methods

A master equation can be used to describe the dynamics on a multi-dimensional PES[254,255], thus giving the time evolution of the probability distribution, $\mathbf{P}(t) = \{P_i(t)\}$, where P_i is the probability of the system being in state i. The master equation has the form

$$\frac{dP_i}{dt} = \sum_j w_{ij} P_j, \tag{5.1}$$

where $w_{ij} = W_{ij} - \delta_{ij} \sum_k W_{ki}$ and W_{ij} is the rate of transition from state j to state i. The diagonal elements of the matrix, w_{ii}, give the total rate of transition out of state i.

If the transition matrix $\mathbf{w}$ is neither decomposable nor splitting, the matrix has a single zero eigenvalue[255], whose eigenvector corresponds to the equilibrium probability distribution, $\mathbf{P}^{eq}$. The master equation describes the evolution of a system from an initial distribution, $\mathbf{P}(0)$, towards $\mathbf{P}^{eq}$; at infinite time the probability distribution must be equal to $\mathbf{P}^{eq}$. The dynamics and thermodynamics must be consistent in this limit.

The transition matrix must satisfy detailed balance to be physically reasonable, i.e. $W_{ij} P_j^{eq} = W_{ji} P_i^{eq}$. As a consequence, the solution of the master equation can be expanded in a complete set of eigenfunctions of the symmetric matrix, $\mathbf{w}'$, defined as $w'_{ij} = \sqrt{P_j^{eq}/P_i^{eq}}\, w_{ij}$. The result is[254]

$$P_i(t) = \sqrt{P_i^{eq}} \sum_{j,k} u_i^j e^{\lambda_j t} u_k^j \frac{P_k(0)}{\sqrt{P_k^{eq}}}, \tag{5.2}$$

where u_i^j is the i^{th} element of the j^{th} eigenvector of $\mathbf{w}'$ and λ_j is the j^{th} eigenvalue of $\mathbf{w}'$. Apart from the zero eigenvalue, the λ_j are all negative. In all our calculations, we have used the above analytical solution. An alternative method of solving the master equation is to integrate equation 5.1 numerically.

The concept of detailed balance also allows us to define when two states come into local equilibrium; i.e. $W_{ij} P_j(t) \approx W_{ji} P_i(t)$. The precise condition that we use is

$$\frac{\left| P_i(t) P_j^{eq} - P_j(t) P_i^{eq} \right|}{\sqrt{P_i(t) P_j(t) P_i^{eq} P_j^{eq}}} \leq 0.01, \tag{5.3}$$

i.e. the two states are within 1% of equilibrium. The equilibration time for two states is then defined as the time at which the above inequality is first satisfied. This allows us to construct equilibration trees mapping out how the system proceeds to the state where the whole system is in equilibrium[256,257].

The standard PES that we consider is depicted in Figure 5.1. It consists of a single funnel in which the number of minima increases rapidly with energy (Figure 5.1b). To simplify the calculations the minima have been grouped into l_{max} levels. The minima in each level are assumed to have identical properties, and to be

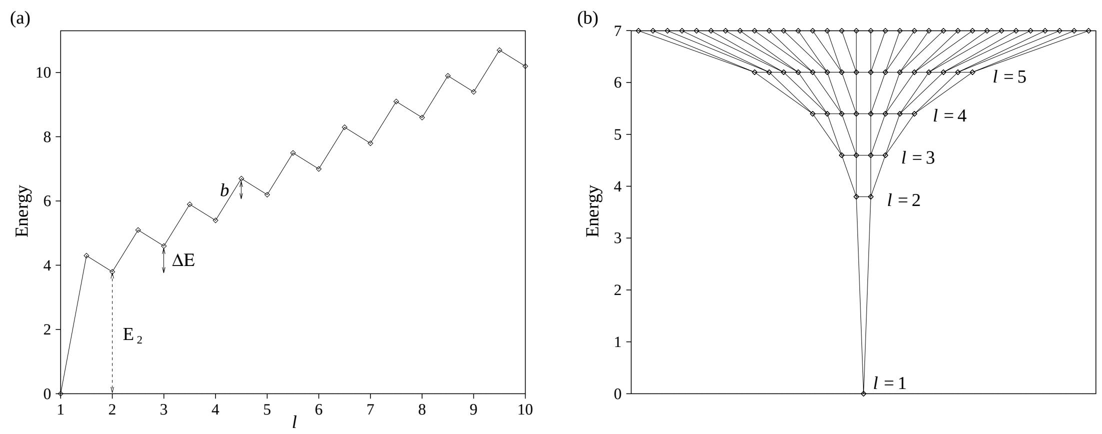

FIGURE 5.1: (a) One-dimensional cross-section through our standard PES showing a reaction pathway from the top of the PES to the global minimum. Integer values of l correspond to minima and half-integer values to transition states. (b) Schematic depiction of a PES with g=2 and l_{max}=6 showing the dramatic increase in the number of minima with energy.

always in equilibrium with each other, thus allowing us to consider each level as a single state in the master equation. This framework is equivalent to assuming that the barriers between minima in the same level are zero. Level $l = 1$ is the global minimum, and the number of minima increases geometrically as the PES is ascended. There are g times more minima in each subsequent level. Therefore the number of minima in level l, n_l, is g^{l-1} and the total number of minima on the PES is $(g^{l_{max}} - 1)/(g - 1)$. We do not consider permutational isomers explicitly since they do not affect the relaxation dynamics. For a real PES, there are $\mathcal{O}(\mathcal{N}!)$ permutational isomers of each minimum. Therefore, there will be many funnels leading down to the different permutational isomers of the global minimum. However, as each funnel is identical they do not need to be considered separately.

We assume that the minima in level l are only connected to minima in levels $l \pm 1$. The transition states connecting minima in adjacent levels are all assumed to lie an energy b above the higher minimum. Each minimum in level l is assumed to be connected to σ minima in level $l-1$ and hence to $g\sigma$ minima in level $l+1$. From this information, and using RRKM theory[166] within the harmonic approximation, the microcanonical rates for transitions between levels are

$$\begin{aligned} W_{l+1,l} &= g\sigma \left(\frac{E - E_{l+1} - b}{E - E_l} \right)^{\kappa-1} \frac{\overline{\nu}_l{}^{\kappa}}{\overline{\nu}_{l+1,l}{}^{\kappa-1}} \quad \text{uphill} \\ W_{l-1,l} &= \sigma \left(\frac{E - E_l - b}{E - E_l} \right)^{\kappa-1} \frac{\overline{\nu}_l{}^{\kappa}}{\overline{\nu}_{l-1,l}{}^{\kappa-1}} \quad \text{downhill,} \end{aligned} \tag{5.4}$$

where E_l is the energy of level l, κ is the number of internal degrees of freedom, and $\overline{\nu}_l$ and $\overline{\nu}_{l,l+1}$ are the geometric mean vibrational frequencies of the minima in level l, and the transition states between levels l and $l+1$, respectively. All other elements of the matrix $\mathbf{W}$ are assumed to be 0 in the present treatment. Similarly, the canonical transition rates are given by

$$\begin{aligned} W_{l+1,l} &= g\sigma \frac{\overline{\nu}_l{}^{\kappa}}{\overline{\nu}_{l+1,l}{}^{\kappa-1}} \exp\left(-\beta(E_{l+1} - E_l + b)\right) \quad \text{uphill} \\ W_{l-1,l} &= \sigma \frac{\overline{\nu}_l{}^{\kappa}}{\overline{\nu}_{l-1,l}{}^{\kappa-1}} \exp\left(-\beta b\right) \quad \text{downhill}, \end{aligned} \tag{5.5}$$

where β is the inverse temperature. We have performed calculations both in the microcanonical and canonical ensembles, but since most of the properties probed show a weak ensemble dependence the majority of the results reported here are microcanonical. The microcanonical ensemble is appropriate to describe an isolated cluster in vacuo, but the canonical ensemble is probably more appropriate for proteins where the solvent can act as a heat bath.

The thermodynamics of the system can be described using a superposition method (Chapter 3), whereby the total energy density of states, $\Omega(E)$, is constructed by summing the density of states for all the energetically accessible minima on the PES. Applying this method to the model PES within the harmonic approximation gives

$$\Omega(E) = \sum_{l=1}^{l_{max}} \frac{n_l (E - E_l)^{\kappa-1}}{\Gamma(\kappa)\overline{\nu}_l{}^{\kappa}}, \tag{5.6}$$

where Γ is the Gamma function. It follows that the microcanonical equilibrium probabilities are given by

$$P_l^{eq}(E) = \frac{n_l (E - E_l)^{\kappa-1}}{\Gamma(\kappa)\overline{\nu}_l{}^{\kappa}\Omega(E)}. \tag{5.7}$$

The canonical partition function and equilibrium probabilities can be constructed in a similar manner:

$$Z(T) = \sum_{l=1}^{l_{max}} \frac{n_l e^{-\beta E_l}}{(\beta h \overline{\nu}_l)^{\kappa}}, \qquad \text{and} \qquad P_l^{eq}(T) = \frac{n_l e^{-\beta E_l}}{(\beta h \overline{\nu}_l)^{\kappa} Z}. \tag{5.8}$$

The global minimum defines the energy zero, i.e. $E_1 = 0$. Beyond $l = 2$ the potential energy is assumed to increase linearly with l. Therefore, $E_l = E_2 + (l-2)\Delta E$ for $l \geq 2$, where ΔE is a measure of the potential energy gradient of the funnel. For a typical cluster PES, the mean vibrational frequency is smaller for minima of higher potential energy[153]—the stabilization of the liquid-like phase at high temperatures is due to both the large number of minima and the greater vibrational entropy. Therefore, we use $\overline{\nu}_l = 1 - (l-1)\Delta\nu$. This defines the unit of time as the vibrational period of the global minimum. The mean vibrational frequency of a transition state is assumed to be the geometric mean of the vibrational frequencies of the two minima it connects, i.e. $\overline{\nu}_{l+1,l} = \sqrt{\nu_l \nu_{l+1}}$

5.3 The Standard Potential Energy Surface

The parameters for the standard PES employed in the present study are

$$\kappa = 101 \qquad \sigma = 1 \tag{5.9}$$

$$E_2 = 3.79688 \qquad \Delta E = 0.8 \tag{5.10}$$

$$g = 10 \qquad l_{max} = 10 \tag{5.11}$$

$$\Delta\nu = 0.01 \qquad b = 0.5 \tag{5.12}$$

These values produce a total of 1.11×10^9 minima on the PES. Some of the thermodynamic properties of the system are shown in Figure 5.2. These were calculated using the following definitions: the microcanonical temperature, T, is given by $1/kT = (\partial \ln \Omega / \partial E)_{N,V}$, the canonical internal energy by $U = -(\partial \ln Z/\partial\beta)_{N,V}$, and the canonical heat capacity by $C_v = (\partial U/\partial T)_{N,V}$. In each case analytic derivatives of the appropriate partition functions were employed.

It can be seen from Figure 5.2 that the system shows the finite-size analogue of a first-order phase transition; there are prominent features in the caloric curves and a large peak in the heat capacity. The transition has a two-state character involving the global minimum and minima in the highest energy region of the PES, with intermediate minima never being significantly populated at equilibrium (Figure 5.2c). The energy at which the global minimum has an equilibrium probability of 0.5, E_f, can be used as a definition of the transition energy. For our 'standard' PES, $E_f = 38.2745$.

Also of note is the Van der Waals loop that occurs in the microcanonical caloric curve (Figure 5.2a), which is a feature unique to finite systems[134]. Such loops have been commonly observed for clusters[135,136] and have also been noted in protein simulations[258]. In small clusters the physical cause of this region of negative heat capacity is the absence of phase separation, which is prevented by the relatively high energetic cost of an interface between the solid-like and liquid-like regions[176].

The microcanonical dynamics of the system were examined by initially populating the highest energy state, i.e. setting $P_{10}(0) = 1$, and observing the relaxation

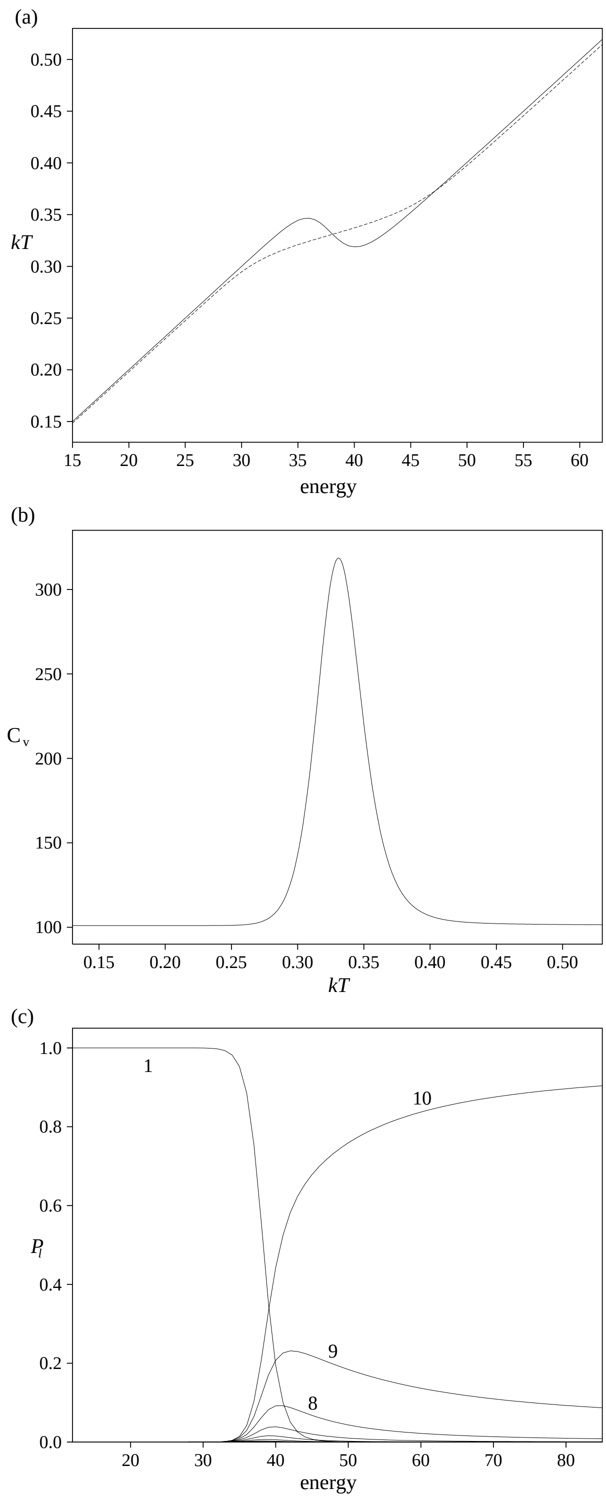

FIGURE 5.2: Thermodynamics of our standard PES: (a) caloric curves for the microcanonical (solid line) and canonical (dashed line) ensembles, (b) canonical heat capacity, and (c) equilibrium probabilities of occupation of level l, P_l^{eq}, for the microcanonical ensemble.

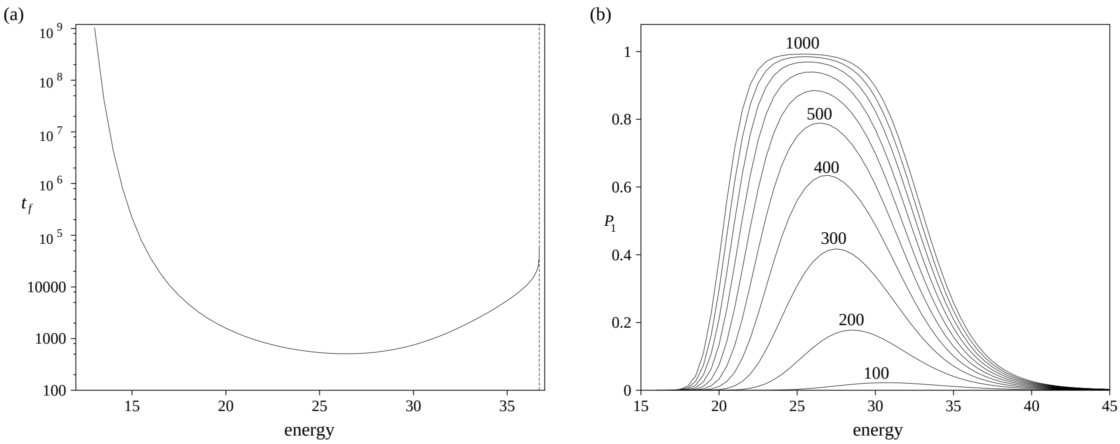

Figure 5.3: (a) Variation of the folding time, t_f, with energy for our standard PES. t_f is defined as the time for P_1 to reach 0.8, from an initial state $P_{10}(0) = 1$. The vertical dashed line denotes the energy at which $P_1^{eq} = 0.8$ and so is the upper limit for which t_f is defined. (Above this energy $P_1(t)$ is always less than 0.8). (b) The variation of the probability of being in the global minimum, P_1, as a function of time and energy. The ten lines shown are for $t = 100$ to 1000 in intervals of 100 and are labelled by the appropriate value of t.

towards equilibrium at constant energy. The corresponding canonical results are omitted for brevity. We consider in particular the time required for the global minimum to develop an 80% probability of occupation. By analogy to relaxation in proteins we call this the 'folding' time, t_f; i.e. $P_1(t_f)$=0.8. The dependence of t_f on E for the standard PES is shown in Figure 5.3a. There is a clear minimum in t_f which occurs at an energy, E_{max}, of 26.41 (= $0.69E_f$). The folding time increases rapidly at low energy because it becomes more difficult to overcome the barriers on the PES, and increases at higher energies because the thermodynamic driving force towards the global minimum is diminished. Similar maxima have been seen in the folding rates of model proteins[247,259–262] in the 'crystallization' rate of a $(KCl)_{32}$ cluster[91], and in the nucleation rate of glasses[263]; these maxima occur for the same reasons as for our model PES. Maxima in the folding rates have also been observed in experiments on proteins which are rapid folders[264–266], but the interpretation of these results is complicated by the fact that real proteins can undergo low temperature denaturation[267].

Another view of the relaxation dynamics is provided by Figure 5.3b which shows the increase in the population of the ground state with time; there is again an energy window for which the global minimum can become significantly populated. It is interesting to note that the energy at which the maximum in the population

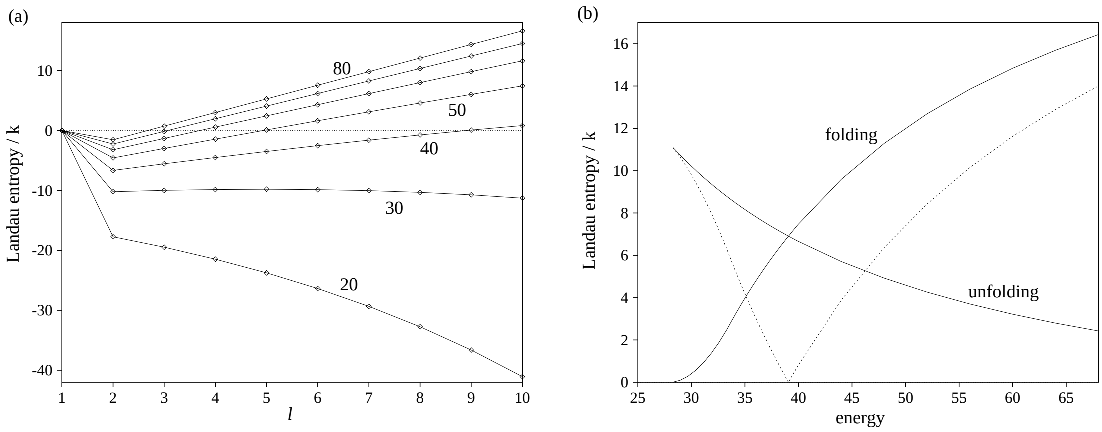

FIGURE 5.4: (a) Landau entropy profiles for our standard PES using the order parameter l. The seven lines shown are for E=20 to 80 in intervals of 10 and are labelled by the appropriate value of E. (b) Depth of the Landau entropy bottleneck relative to the Landau entropy maximum at large values of l ('folding') and relative to the Landau entropy maximum at the global minimum ('unfolding'). The dashed line gives the magnitude of the Landau entropy difference between the two stable states.

of the global minimum occurs decreases somewhat as the time increases. This indicates that to optimize the population of the global minimum one should not quench the system to a low energy and let it evolve, but rather use a cooling schedule whereby the energy is decreased as a function of time, as in constant thermodynamic speed annealing[268].

To further understand the dynamics of the system it is helpful to use Landau functions. In the microcanonical ensemble the presence of a first-order phase transition is indicated by two maxima in the Landau entropy separated by a well. We choose to call the minimum in the Landau entropy an entropy bottleneck, since it represents a constricted region of phase space which has a lower density of states than the regions that it connects. The Landau profiles also enable us to understand the influence of the thermodynamics on the dynamics. If there is one stable state, relaxation to that state is likely to be rapid unless other kinetic factors intervene. If there are two stable states, the rate of transition from the metastable to the stable state decreases as the entropy bottleneck becomes narrower or the free energy barrier increases. The importance of the Landau free energy in understanding protein folding has been recognized by Bryngelson *et al.*, who considered the dynamics expected for a number of different Landau free energy profiles[245].

Application of a Landau-type analysis using a restricted thermodynamic potential depends on finding a suitable order parameter which can differentiate between the states of interest. For many systems this can be a difficult task, but for our model PES l provides a natural order parameter since it is a measure of the distance from the global minimum. The Landau entropy is then defined as $S_L(l) = S(E) + k \ln P_l^{eq}(E)$ and the Landau free energy by $A_L(l) = A(T) - kT \ln P_l^{eq}(T)$. Landau entropy profiles for our standard PES are given in Figure 5.4 for a range of energies. At low energy there is a single maximum corresponding to the global minimum, and relaxation to the global minimum is thermodynamically 'downhill'. Consequently at E_{max}, the probability density 'flows' smoothly down the PES with the intermediate states having significant transient populations (Figure 5.5).

The equilibration trees in Figure 5.5 map out the times at which different levels come in to approximate equilibrium with each other (as defined by equation 5.3). At E_{max} levels 1 and 2 first reach local equilibrium (Figure 5.5e), i.e. equilibrium first occurs between the levels corresponding to the Landau entropy maximum. Next, level 3 reaches equilibrium with levels 1 and 2, and so on to higher values of l: the local equilibrium propagates to higher energy away from the Landau entropy maximum. Similar behaviour occurs at lower energies, except that the barriers on the PES have a larger retarding effect on the dynamics. Therefore, the equilibration tree at E=15 (Figure 5.5c) shows the same structure as at E_{max} except that the time scales are around a thousand times slower. In the microcanonical ensemble the kinetic energy increases as the potential energy decreases. Consequently, at low energies the transition rate over the barriers is much slower at the top of the PES, and the system accelerates as it descends the PES. Therefore, transient populations in the intermediate states are less pronounced at $E = 15$ than at E_{max}. In the canonical ensemble this effect is absent because the kinetic energy is independent of the position in configuration space.

For $E > 28.238$, there are two Landau entropy maxima, which correspond to the global minimum and the high potential energy 'unfolded' states. The entropy bottleneck occurs because at lower potential energy the loss in entropy due to the reduced number of minima is not completely compensated by the greater vibrational entropy that results from the larger kinetic energy. As a result of the entropy bottleneck, the equilibration trees have a qualitatively different form from those at lower energies; equilibration initially occurs within the secondary Landau entropy maximum corresponding to the higher lying levels and then propagates down the PES (Figure 5.5f). It is interesting to note that this change in direction of equilibration appears at approximately the same energy as the Landau entropy bottleneck. For E=36 there is a separation of time scales between equilibration

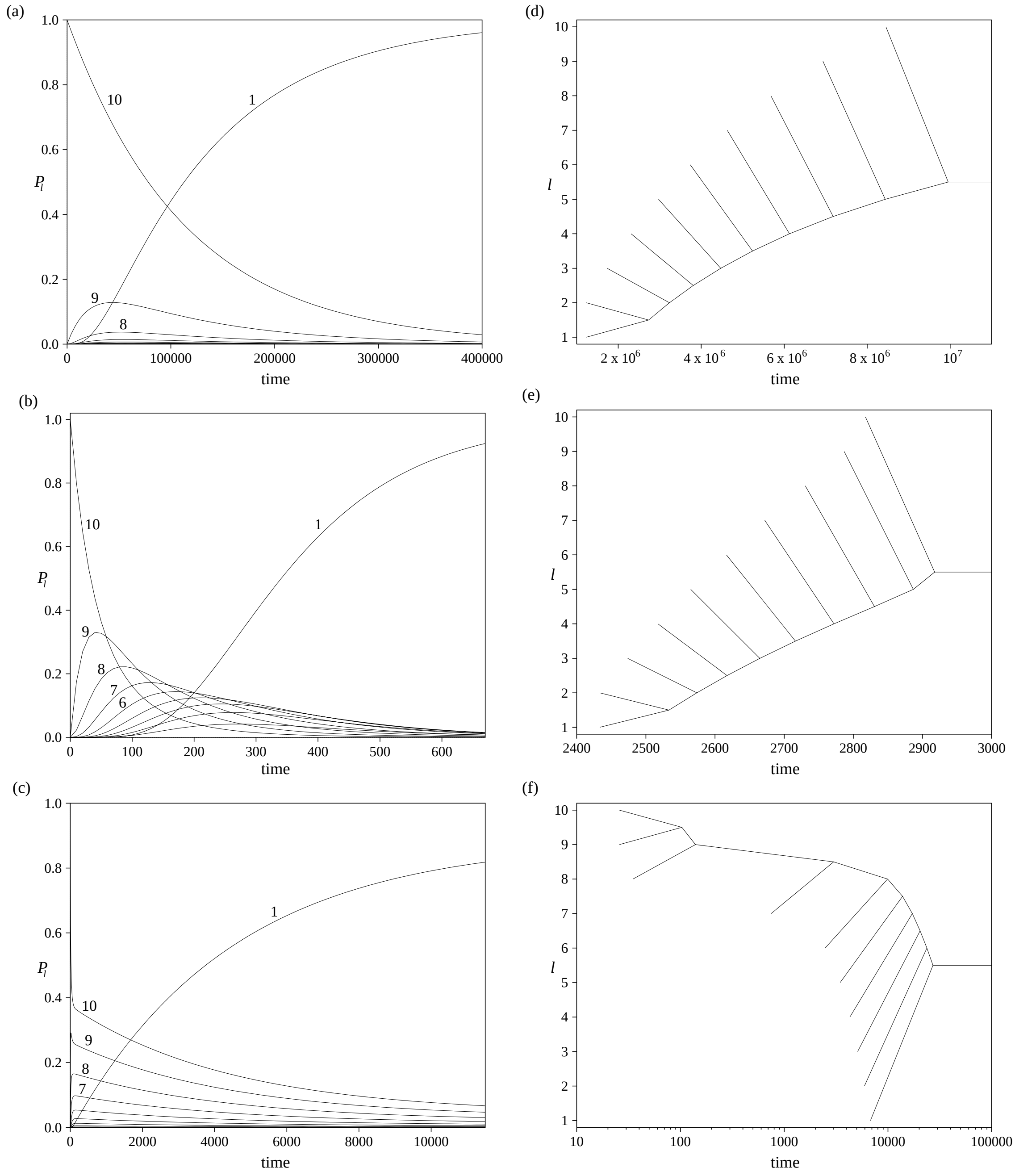

FIGURE 5.5: (a)–(c) Time evolution of the probabilities, P_l, and (d)–(f) equilibration trees for our standard PES at an energy of (a) and (d) 15; (b) and (e) 26.41; and (c) and (f) 36. Initially, P_{10}=1. The energy 26.41 corresponds to the minimum in the 'folding' time of Figure 5.3a. Each line in the equilibration tree represents a single level or a set of levels that are in equilibrium with each other. The lines join at times when states first come into local equilibrium as defined by equation 5.33. The vertical position of these nodes is given by the average value of l for the levels that are in equilibrium with each other. The horizontal positions of the ends of lines on the left-hand side of the graph, which correspond to individual levels, are arbitrary.

within the secondary Landau entropy maximum, and equilibration between the maxima. This is reflected in the time evolution of the probability distribution: the probabilities of occupying higher potential energy levels rapidly attain approximate local equilibrium values which then slowly decay as the global minimum becomes more populated (Figure 5.5c).

Similar behaviour is seen for the Landau free energy in the canonical ensemble. In fact, the temperature at which A_L develops a double well can be derived analytically (assuming that $\Delta\nu$ is small). Two free energy minima are observed if

$$kT > \frac{\Delta E}{\ln g + \kappa \Delta\nu}. \tag{5.13}$$

This corresponds to $kT > 0.242$ for our standard PES. As relaxation to the global minimum is impeded by a free energy barrier, equation 5.13 implies that large ΔE and small g and $\Delta\nu$ assist relaxation to the global minimum. These are the same conditions as those that lead to a large value of T_f.

Landau free energy profiles have previously been calculated for lattice models of proteins[233,234,247,250,258]. They show a similar behaviour to that found for our model PES, with a single minimum at low temperature, and two stable states at higher temperatures corresponding to the folded and unfolded protein[233,234,247]. For clusters there can be a range of temperature around the melting transition for which Landau functions reveal two stable states corresponding to the solid-like and liquid-like states[112,133]. Similarly it has been shown that the nucleation of a crystalline phase from the bulk Lennard-Jones liquid involves a large free energy barrier[269,270].

5.4 Modified Potential Energy Surfaces

The previous section showed that our simplest model PES captured many of the dynamical effects commonly seen in finite systems such as proteins and clusters. Here we seek to modify the standard PES to understand which factors influence a system's ability to find the global minimum. The model PES has too many parameters to allow a systematic study of the whole parameter space. Additional complexity could, of course, be generated by making the parameters functions of l. The effects of varying some of the parameters, though, are fairly obvious. For example, increasing σ would make all the rates σ times faster. Our approach here is to make only changes to the PES that leave the basic thermodynamics

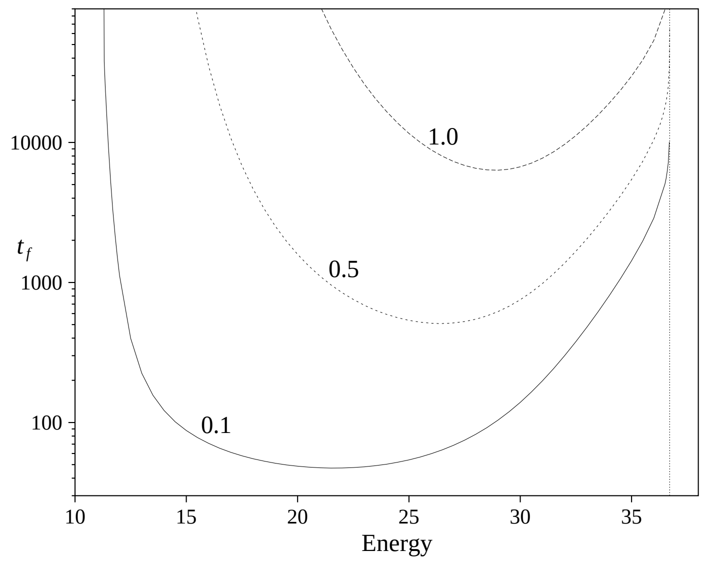

Figure 5.6: The dependence of t_f on the barrier height, b, for our standard PES. t_f is defined by $P_1(t_f)$=0.8. The vertical dotted line denotes the energy at which $P_1^{eq} = 0.8$. Each line is marked by the value of b.

unchanged—i.e. we keep E_f constant—and to concentrate on the factors that could lead to particularly interesting behaviour.

5.4.1 Barrier heights

The first modification that we make is to change the barrier height. Unsurprisingly, increasing b leads to a reduction in the rate of relaxation to the global minimum (Figure 5.6). More interesting is the dependence of this effect upon the energy: increasing the barrier heights leads to a much larger increase in t_f at low energies. This is because a larger fraction of the kinetic energy must be converted to potential energy to overcome the barriers at low energies. The energy at which the maximum 'folding' rate occurs increases as the barrier heights increase. The shape of the minimum in t_f is also dependent on b, becoming broader as the barrier heights are decreased; at $b = 0.01$ there is a larger energy window for which fast relaxation to the global minimum can occur.

Similar behaviour is observed for the canonical ensemble, namely the rate of relaxation to the global minimum has a $\exp(-\beta b)$ dependence on the barrier height. Therefore, at fixed temperature the relaxation rate has a rigorously exponential dependence on b, and changing b has a greater effect on the relaxation rate at low temperature.

These results can easily be understood in the language of Bryngelson and

Wolynes[243–245]: increasing the barrier heights corresponds to increasing T_g without changing T_f, hence decreasing the ratio T_f/T_g and the 'folding' ability. Potassium chloride clusters provide a good example of systems for which relatively small barrier heights produce rapid relaxation down the PES[252].

5.4.2 The slope of the funnel

In the energy landscape model of protein folding, the slope of the funnel (potential energy gradient) is a key parameter[246,250]. Here we analyse the effect of the slope by varying ΔE. To keep E_f constant we must also vary E_2. For the case when ΔE is zero, i.e. a flat surface, $E_l = 10$ for $l \geq 2$. It can be seen from Figure 5.7a that the slope has a dramatic effect on the folding time, and that a sufficiently steep slope is necessary for reasonably rapid relaxation to the global minimum. This behaviour can be explained from the relative rates of uphill and downhill transitions in our model. The barrier for going downhill is b, and for going uphill $\Delta E + b$. Therefore, the downhill transition rate is not significantly affected by the slope, but the uphill rate decreases rapidly as the slope is increased. For the flat PES taking a step to higher l is, ignoring the vibrational contributions to the rate, g times more likely than to lower l, causing the probability distribution to remain concentrated at large values of l. As a slope is introduced into the PES the uphill rate decreases relative to the downhill rate and the folding time begins to decrease rapidly. However, once the PES is sufficiently steep to make the uphill rate small compared to the downhill rate, further increasing the slope only has a limited effect. At large ΔE the folding time even begins to increase again because the level $l = 2$ begins to have a low enough energy for it to have a significant equilibrium population.

An equivalent explanation for the behaviour can be provided in terms of the Landau entropy. In order to keep E_f constant as the slope is decreased the energy of the intermediate levels must be increased. This has the effect of reducing the equilibrium probability for occupation of these levels. Therefore, as ΔE is decreased a Landau entropy bottleneck develops and then increases in depth (Figure 5.7b). The folding time increases rapidly as the bottleneck deepens, but has only a weak dependence on ΔE when there is a single Landau entropy maximum.

The flat PES would correspond to the Levinthal limit where the states are searched randomly, but for the variation of the vibrational frequency with l, which makes the minima in higher levels slightly more probable. For this PES, equili-

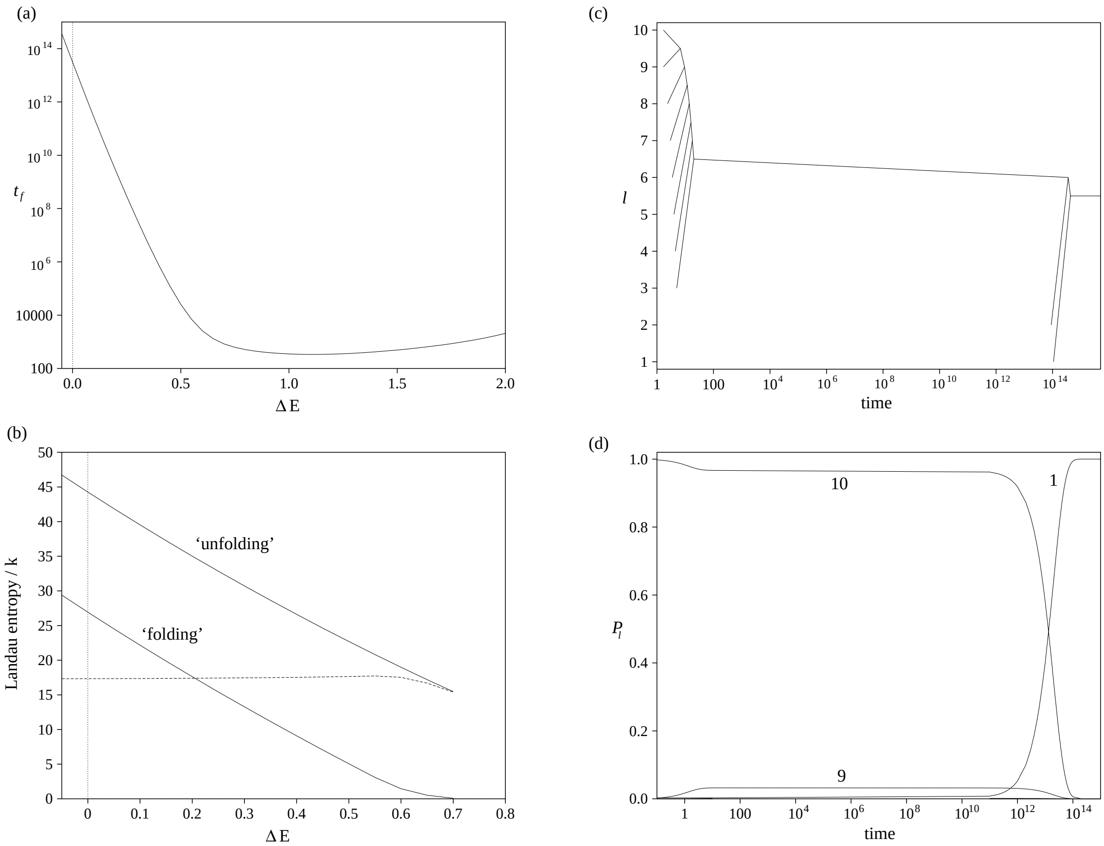

FIGURE 5.7: (a) Variation of t_f with the slope of the PES. (b) Variation of the magnitude of the Landau entropy bottlenecks (solid lines) with ΔE. The dashed line gives the magnitude of the Landau entropy difference between the two stable states. (c) Equilibration tree and (d) time evolution of the probabilities, P_l, for the flat PES (ΔE=0) from an initial distribution P_{10}=1. All results are for E=26.41.

bration occurs initially at large values of l and proceeds away from the Landau entropy maximum. There is a large separation (of the order of 10^{13}) between the time scales for equilibration within the large l entropy maximum and between the entropy maxima (Figure 5.7c), because of the deep Landau entropy bottleneck. This is evidenced by the probability flows across the PES (Figure 5.7d); P_{10} and P_9 rapidly reach local equilibrium values, which then decay away as the ground state is slowly populated.

These results help to explain why potassium chloride clusters are able to relax to the rock-salt structures so rapidly[91]. Reaction pathways leading down to the rock-salt structures generally exhibit[252] $\Delta E \gg b$.

As ΔE is decreased E_2, which is simply the energy gap between the global minimum and the next lowest energy state, has to be increased to maintain E_f. Therefore, the efficiency with which the global minimum is found is reduced as

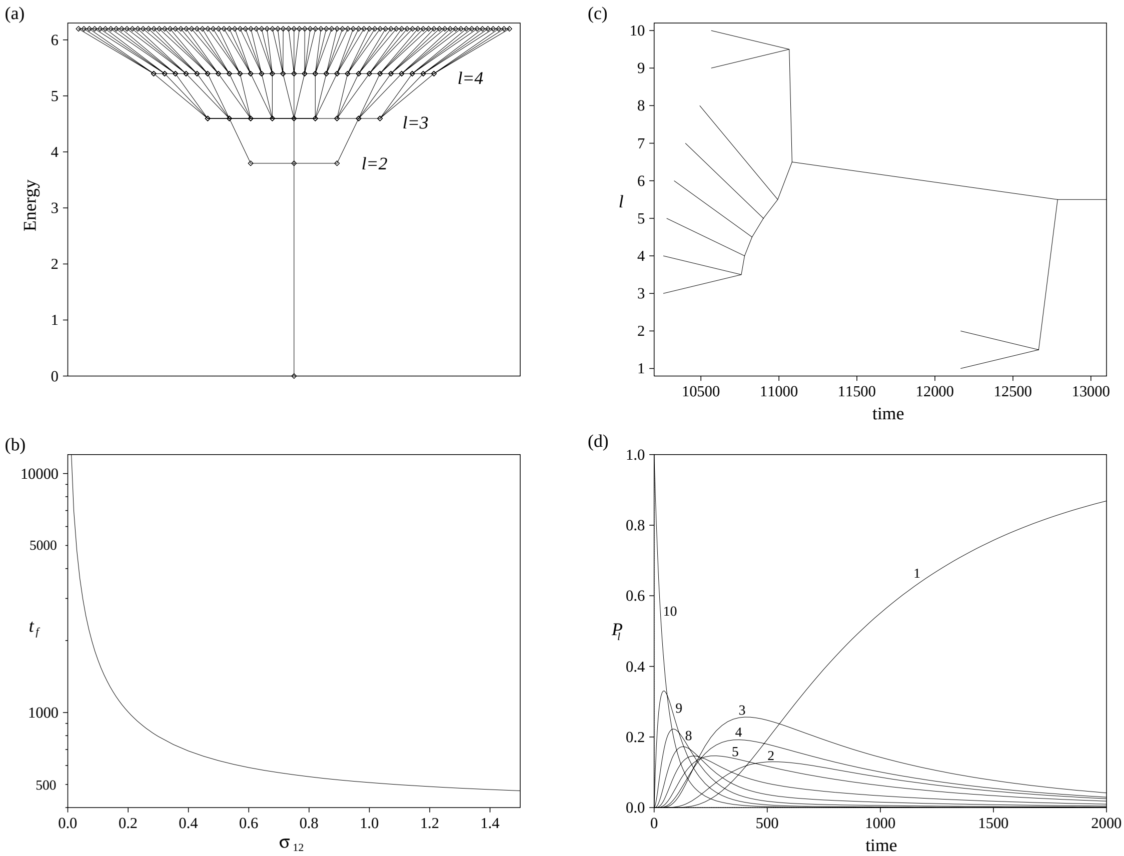

FIGURE 5.8: (a) Schematic depiction of a PES with a kinetic bottleneck. $g = 3$, and $\sigma = 1$ except for $\sigma_{12} = \sigma_{23} = \frac{1}{3}$. (b) Dependence of t_f on σ_{12}. (c) Equilibration tree and (d) time evolution of the probabilities, P_l, from an initial distribution P_{10}=1 for a PES with $\sigma_{12} = 0.1$. All results are for E=26.41.

E_2 is increased. This result does not necessarily contradict the view of Sali *et al.*, however, since we can decouple factors in our model in a way that may not be possible for real PES's. However, it does show that even if the energy gap view were correct, it would require further explanation; it is not self-evident why an energy gap should guarantee kinetic foldability.

5.4.3 Kinetic bottlenecks

For our standard PES surface σ is independent of l. However, a real PES is unlikely to show such uniformity. For example, it has been suggested that for some proteins the number of available pathways may decrease near to the native state leading to the observation of kinetically-determined intermediates[245].

In this section we investigate the effect of making σ dependent on l. In particular, we reduce the number of connections between l=1 and l=2, and between

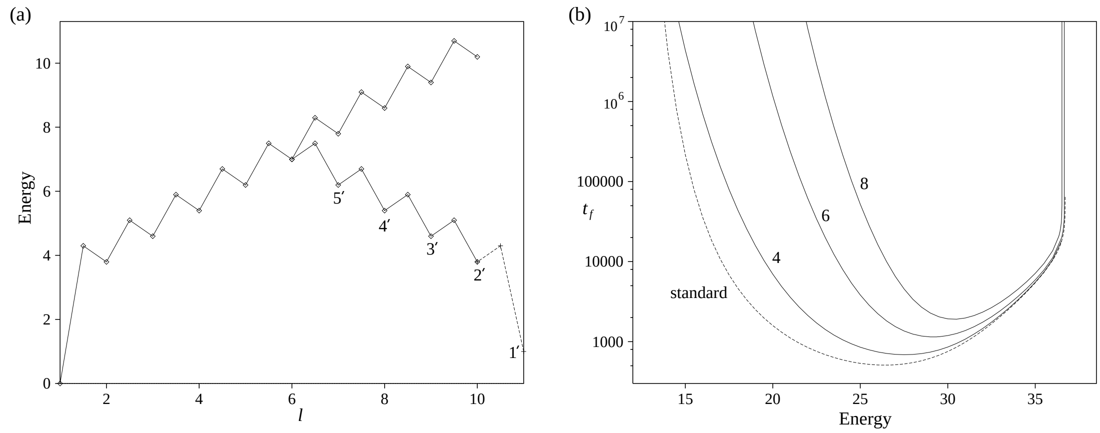

FIGURE 5.9: (a) One-dimensional cross-section through a PES with two funnels where l_{node}=6 The values of l' for the minima in the secondary funnel are as labelled. (b) Dependence of t_f upon the energy and l_{node}. The values of l_{node} are as labelled, and the curve for our standard PES (dashed line) has been added for comparison.

l=2 and l=3. Such a PES is illustrated in Figure 5.8a; it has $\sigma_{12} = \sigma_{23} = \frac{1}{3}$ and so only 1 in 3 of the minima in level 3 are connected to a minimum in level 2. These changes do not affect the thermodynamics in any way (including the Landau entropy) but instead produce a kinetic bottleneck. In Figure 5.8b, the dependence of the folding time for the modified surface upon σ_{12} is shown. t_f increases as the number of connections is reduced, but at the most by only one to two orders of magnitude. The slowing down of the dynamics at the bottleneck leads first to local equilibrium above the bottleneck (Figure 5.8c) and to a build-up of probability in level 3 and to a lesser extent in levels 4 and 5 (Figure 5.8d).

These results show that kinetic bottlenecks may significantly affect the folding rate, but to a smaller extent than other factors, such as the slope of the funnel or the multiple funnels that we shall consider next. Kinetic bottlenecks may be important in rationalizing the detailed dynamic behaviour of a particular system, but are unlikely to represent the difference between, for example, folding and non-folding proteins.

5.4.4 Multiple funnels

In this section, we consider the consequences for the dynamics when the PES has more than one funnel. This allows us to examine Leopold *et al.*'s suggestion that the amino acid sequences which fold to a unique native state are characterized by a single funnel leading to the global minimum, whereas those that do not fold have multiple funnels leading to structurally different minima[248]. The results should also

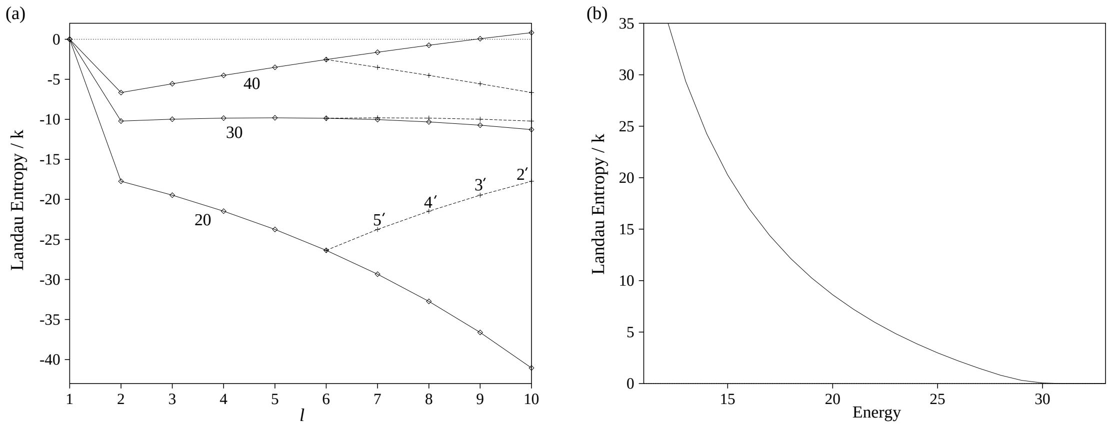

FIGURE 5.10: (a) Landau entropy profiles for a PES with two funnels and l_{node}=6. The lines are labelled by the corresponding values of the energy. The dashed lines are for the secondary funnel and points have been marked by the appropriate value of l'. (b) The energy dependence of the magnitude of the Landau entropy bottleneck for passage from the bottom of the secondary funnel to the primary funnel.

be relevant to other systems where the system has difficulty in reaching the global minimum, such as glasses. It has been suggested that the glassy states are located at the bottom of funnels which are separate from the crystalline funnel[91,271,272].

A schematic picture of the modified PES which we have investigated is shown in Figure 5.9. It has a secondary funnel joined to the primary funnel at $l = l_{node}$. The levels in the secondary funnel are denoted by l' and have the same properties as the corresponding levels for the standard PES. Initially, we consider the case where the secondary funnel ends at $l' = 2'$, since for this case the secondary funnel hardly affects the equilibrium thermodynamics—the $P_{l'}^{eq}$ never have significant magnitudes (see e.g. Figure 5.2c).

The rate of relaxation to the global minimum is not significantly affected by the secondary funnel at energies near to E_f, but as the energy is decreased the rate slows dramatically compared to the standard PES (Figure 5.9b). The latter effect becomes more pronounced at larger values of l_{node}. This behaviour can be explained by examining the Landau entropy (Figure 5.10). At higher energies, the bottom of the secondary funnel is not an entropy maximum, and so the system never gets trapped there. However, below the energy for which relaxation down the primary funnel is thermodynamically favourable, level $l' = 2'$ becomes a Landau entropy maximum and there is a Landau entropy bottleneck which must be overcome in passing from the secondary to the primary funnel (Figure 5.10b). The conditions which are thermodynamically most favourable for folding for the standard PES

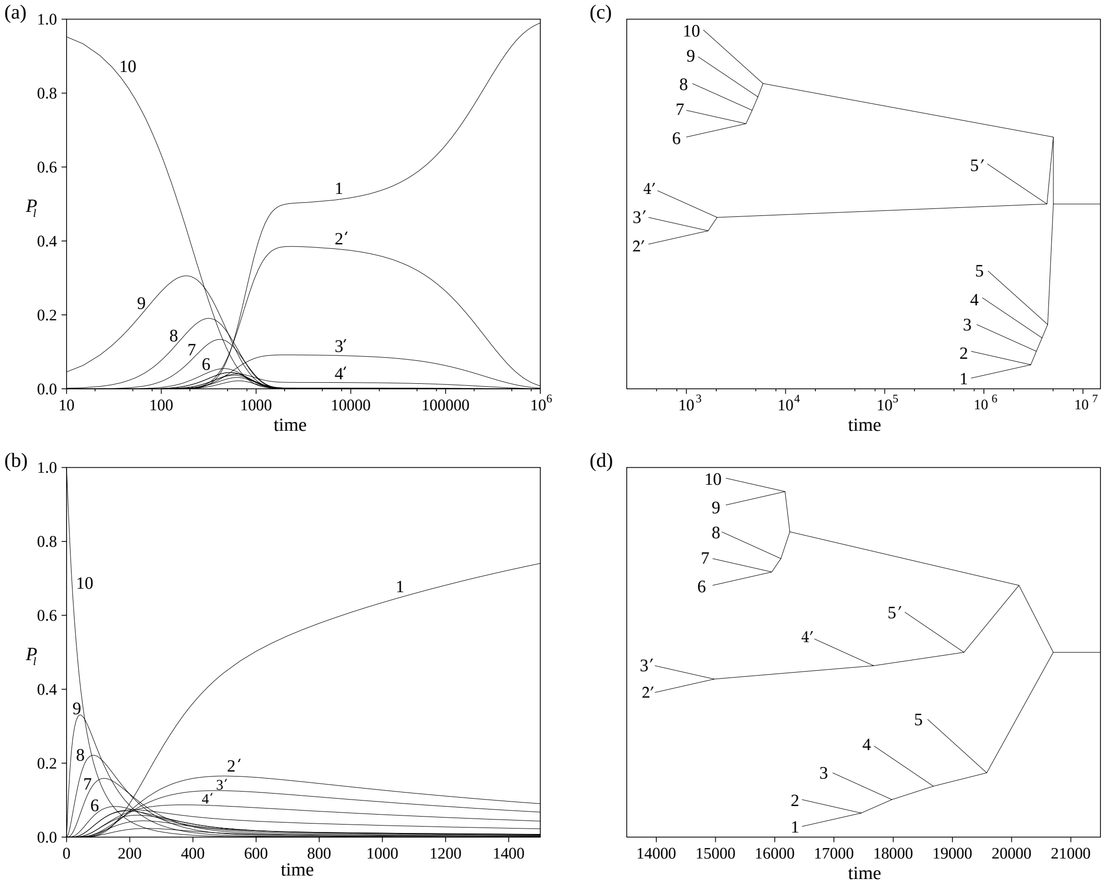

FIGURE 5.11: (a) and (b) Time evolution of the probabilities, P_l, and (c) and (d) equilibration trees at an energy of (a) and (c) 21; and (b) and (d) 26.41 for a PES with two funnels and l_{node}=6. The secondary funnel ends at l'=2. Initially, P_{10}=1. The lines of the equilibration trees corresponding to individual levels are labelled by the appropriate values of l or l'.

also most favour trapping in the secondary funnel.

The consequences of these changes in the Landau entropy can be seen in the probability flows and the equilibration trees (Figure 5.11). At $E = 26.41$, there is only a small entropy bottleneck associated with exit from the secondary funnel, and the time scale for this process is only about an order of magnitude larger than that for relaxation from high energy minima. At higher energies, where there is no entropy bottleneck, this time scale separation disappears altogether. However, for $E = 21$ there is a much larger entropy bottleneck to be overcome. By $t = 1000$, the system has relaxed down the PES, reaching a state where $P_1 \approx \frac{1}{2}$ and the probability of being in the secondary funnel is also close to a half (Figure 5.11a). In the initial relaxation the system is equally likely to go into either funnel, because both funnels have the same properties around $l = l_{node}$. This situation persists for about two orders of magnitude in time (note the log scale in the figure) until the probability begins to flow from the secondary funnel into the primary. The same basic structure is seen for both equilibration trees in Figure 5.11; equilibration occurs separately above l_{node} and within the bottoms of the two funnels. The tree for $E = 21$ clearly shows the large separation in time scales.

There have been some interesting investigations of the effects of trapping by low energy non-native states in protein folding simulations[262,273]. These show that at low temperatures there is a fast component to the relaxation rate, which corresponds to the protein passing directly into the native structure, and a slower component, where the system becomes trapped in one of the low energy non-native states before reaching the native structure. As for our model PES, this separation of time scales can disappear at higher temperatures[262]. Experimental studies of protein folding often show kinetics with two (or more) time scales present[264,274–278]; there are subpopulations of proteins which fold at different rates. Experiments on hen lysozyme have clearly shown that the faster rate corresponds to direct folding to the native state[276], and it has been inferred that the slower rate is due to trapping in low energy non-native structures. This interpretation has been confirmed for cytochrome c. For this protein, the chain misorganization responsible for trapping is the incorrect ligation of a haem group by a histidine residue. Under refolding conditions which prevent this incorrect contact forming, the protein folds on a time scale of 15ms, but when trapping occurs folding takes longer than 0.3s[278]. The temperature dependence of the two rate constants[264] for cytochrome c agrees well with the behaviour of our model PES. The faster rate constant exhibits a maximum, as would be expected for direct relaxation to the global minimum (Figure 5.3a), whereas the rate constant for the slower process just decreases as the temperature decreases[264], which is what would be expected when the magnitude of the Landau entropy bottleneck for escape from a trap increases with decreasing temperature (Figure 5.10b).

Interestingly, it has been suggested that chaperonins, such as GroEL and GroES, assist protein folding by reducing the effect of trapping[279,280]. The chaperonins bind proteins which are trapped in non-native structures and then release them in a less folded state, thus allowing the protein another opportunity to find the native structure directly. Simulations show that the folding rate of slow-folding proteins increases significantly when this 'iterative annealing' model for chaperonin action is included[281], lending further credence to this proposal.

The trapping properties of the secondary funnel can be increased by adding

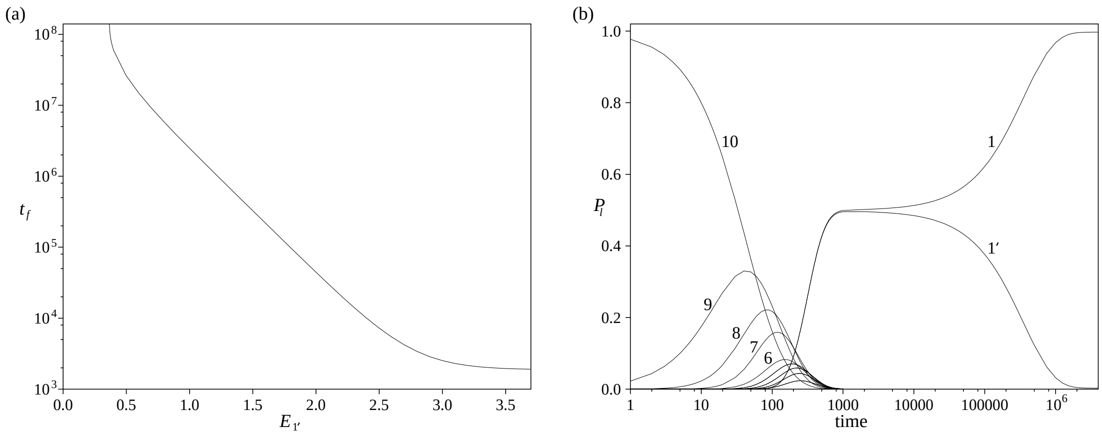

FIGURE 5.12: Dynamic properties of a PES with two funnels and l_{node}=6 at an energy of 26.41. The secondary funnel ends at $l' = 1$. (a) Dependence of t_f upon $E_{1'}$. (b) Time evolution of the probabilities, P_l, for $E_{1'}$=1.5 from an initial distribution P_{10}=1.

a minimum at $l' = 1'$, as shown in Figure 5.12a. t_f increases exponentially as the depth of the secondary funnel is increased. As $E_{1'}$ approaches the value for which $P_{1'}^{eq}$=0.2 ($E_{1'}$=0.364), t_f increases even more rapidly, as expected from our definition of t_f ($P_1(t_f) = 0.8$). At higher energies, the dependence of the relaxation rate on $E_{1'}$ diminishes as $E_{2'}$ is approached. For this system the probability density flows rapidly and equally into the bottoms of both funnels and then begins to trickle very slowly back up the PES from $l' = 1'$ into the primary funnel (Figure 5.12b). This PES provides an example where the energy difference between the two lowest energy minima has a very significant effect on the relaxation rate to the global minimum, and supports our suggestion that this quantity is most important when the two minima equilibrate very slowly because they are structurally dissimilar and well-separated on the PES.

The PES of $(KCl)_{32}$ shows some interesting features. The relaxation down to the low energy rock-salt-like structures from the amorphous states is rapid[91]. However, it is not clear whether the system can then reach the $(4 \times 4 \times 4)$ global minimum from the other rock-salt-like structures, which can have very different shapes from the global minimum, e.g. the second lowest energy minimum has a $(5 \times 4 \times 3)$ shape plus an extra row of four atoms[127]. It might be that these crystal-like structures have separate funnels, the steep slopes of which make interconversion to other rock-salt-like structures difficult. This would give rise to a significant separation between time scales for relaxation down the PES and equilibration between the rock-salt-like structures, similar to that seen for our model PES above.

In the above examples, the properties of both funnels were the same except at $l' = 1'$. As a result, the initial relaxation down the PES was equally likely to take the system into either funnel. However, for materials that easily form glasses the system is more likely to relax into the amorphous glassy minima than into the crystalline minima. A similar effect can be achieved in our model by making the 'width' of the primary funnel narrower than that of the secondary. g determines the number of minima in each level, and consequently also determines the increase in the configurational entropy as the funnel is ascended, i.e. its width. Therefore, if we make the value of g for the primary funnel below l_{node}, g_p, where g_p is smaller than for the rest of the PES, the primary funnel is only connected to $g_p^{l_{node}-1}$ of the $g^{l_{node}-1}$ minima in the level $l = l_{node}$. For the parameters l_{node}=6, g_p=4 and g=10, this corresponds to only 1% of the minima in the level $l = l_{node}$. These changes to the PES affect the rate of transition between levels l_{node} and $l_{node} - 1$, which becomes

$$W_{l_{node}-1,l_{node}} = \sigma \left(\frac{g_p}{g}\right)^{l_{node}-1} \left(\frac{E - E_{l_{node}} - b}{E - E_{l_{node}}}\right)^{\kappa-1} \frac{\overline{\nu}_{l_{node}}{}^{\kappa}}{\overline{\nu}_{l_{node}-1,l_{node}}{}^{\kappa-1}}. \qquad (5.14)$$

This PES can mimic some of the effects of frustration. If, as in simple liquids, the high potential energy amorphous minima have a different type of order from the crystalline global minimum, one might expect that the glassy states at the bottom of the liquid-like band of minima maximize the polytetrahedral order as far as is possible. Therefore, the glassy minima are not likely to be connected to the crystalline funnel, but instead reside at the bottom of rugged funnels which descend from the liquid-like regions of configuration space. In contrast, the funnel terminating at the crystalline state is likely to be connected to a small fraction of the higher energy amorphous minima which contain incipient crystal-like nuclei. Consistent with this idea, large Landau free energy barriers have been found in simulations of nucleation of a crystalline phase from a Lennard-Jones liquid[269,270]. This view of the PES topography of glass-formers has many similarities to the ideas put forward by Stillinger for fragile liquids[271,272].

We now choose to end the secondary funnel at $l' = 2'$ since glasses normally have a significantly higher potential energy than the crystalline state. For this PES, the rate of relaxation to the global minimum dramatically slows as the primary funnel is narrowed by decreasing g_p (Figure 5.13a). This occurs because the level l=5 has a lower entropy than $l' = 5'$ and $l_{node} = 6$ (Figure 5.13b). On reaching $l = l_{node}$, the system is much more likely to pass into the secondary funnel, rather than overcome the Landau entropy bottleneck associated with passage into the

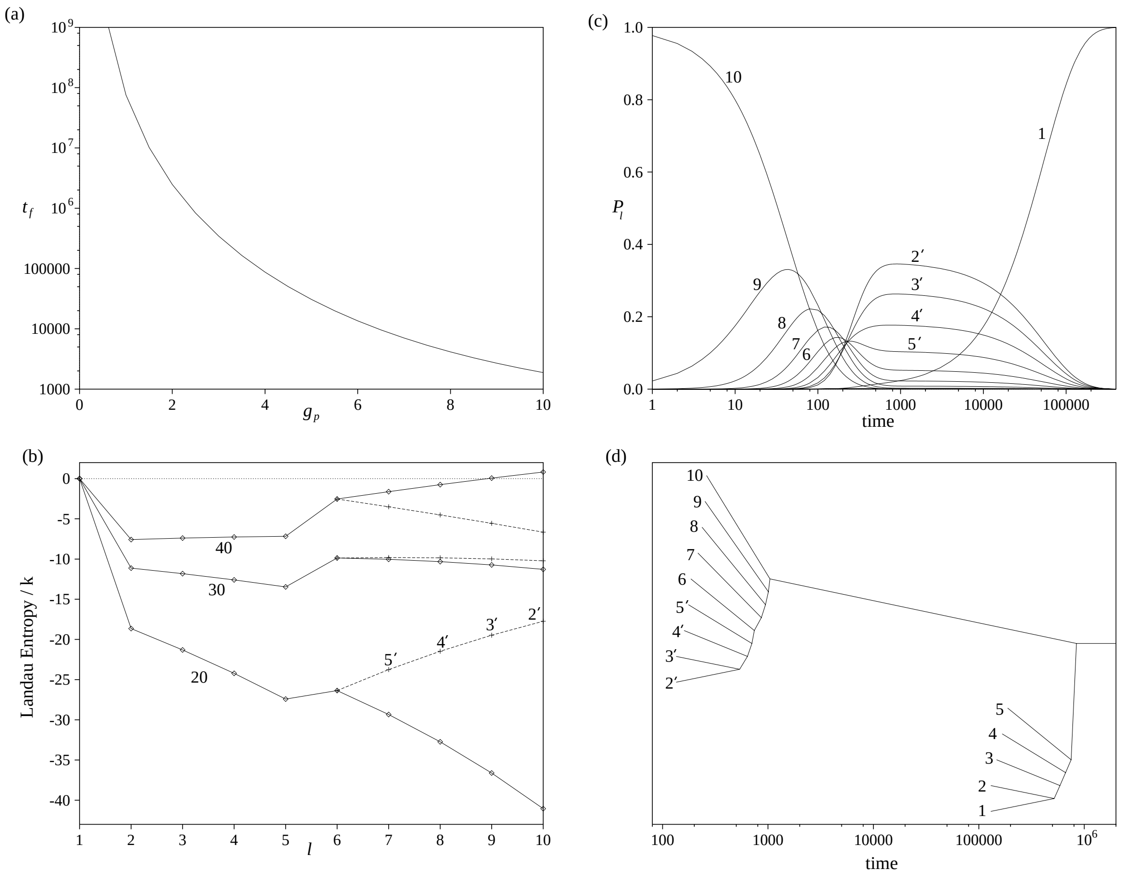

FIGURE 5.13: Dynamic properties of a PES with two funnels and l_{node}=6. For $l < l_{node}$ the primary funnel has a lower value of g than for the rest of the PES. The secondary funnel ends at $l' = 2$. (a) Dependence of t_f upon g_p at an energy of 26.41. (b) Landau entropy profiles when g_p=4 at different values of the energy as labelled. (c) Time evolution of the probabilities, P_l, and (d) Equilibration tree when g_p=4 from an initial distribution P_{10}=1 at an energy of 26.41.

primary funnel. The depth of this bottleneck increases as the width of the primary funnel decreases.

The time evolution of the probabilities when g_p=4 shows that all the probability initially passes into the secondary funnel, and then slowly trickles back through the Landau entropy bottleneck into the global minimum (Figure 5.13c). This results in a large separation of time scales between the time for equilibration within the secondary funnel and between the two funnels (Figure 5.13d).

This example shows that the secondary funnel is a more effective trap when it has a larger width than the primary funnel. Earlier in this section, we showed that a low energy minimum at the bottom of the secondary funnel also increases its trapping efficiency. By combining these two features we can create a PES where

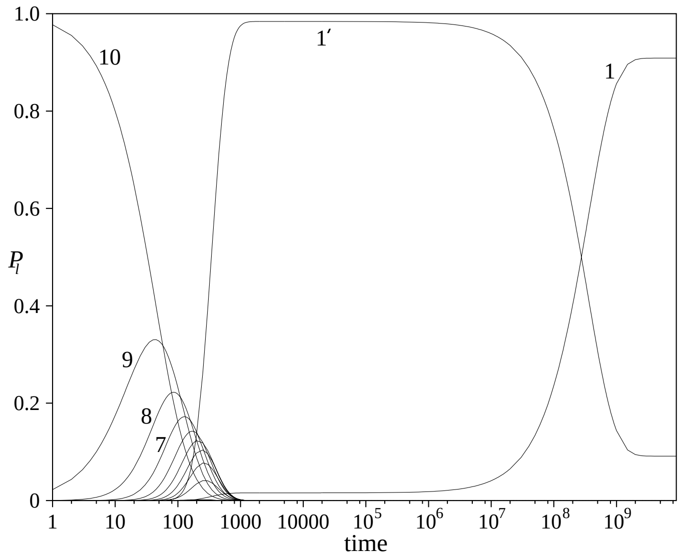

FIGURE 5.14: Time evolution of the probabilities, P_l, for a PES with two funnels and l_{node}=6 at a total energy of 26.41. For $l < l_{node}$ the primary funnel has g_p=4. The secondary funnel ends at l'=1 at a potential energy $E_{1'}$=0.6. Initially, P_{10}=1.

virtually all the probability flows to the bottom of the secondary funnel, and then remains there for a long time (Figure 5.14). Only on very long time scales does the global minimum start to become significantly populated; t_f for this PES is 7.4×10^8. In the example shown $E_{1'}$=0.6, and so $P_{1'}^{eq}$=0.091, thus preventing complete population of the global minimum at the total energy considered. This PES provides an example of how relaxation to a unique metastable minimum might occur, simply because of its greater kinetic accessibility. A similar behaviour is seen for some proteins from the serpin family of protease inhibitors[282]; the protein plasminogen activator inhibitor-1 first folds to the active state, but then on a time scale of hours can spontaneously transform to an inactive latent form[154,283].

The above PES bears some resemblance to the 38-atom and 75-atom Lennard-Jones clusters, for which the low energy icosahedral minima are reached on relaxation down the PES, rather than the fcc or decahedral global minimum[73,89]. The difficulty in reaching the global minima is associated with the narrowness of the funnels leading down to the global minima. Our results also suggest that it would not be so surprising if it was effectively impossible to reach the global minimum of the 75-atom cluster by a dynamic method without biasing the system towards decahedral minima. These clusters, though, have an additional complication compared to our model PES. The free energy global minimum depends on temperature, and only at low temperatures does it actually correspond to the global potential energy minimum. This is because of the greater configurational entropy of the minima at the bottom of the icosahedral funnel (§2.4).

5.5 Summary

In this chapter we have analysed model PES's of increasing complexity using a master equation approach. The existence of a suitable order parameter for these surfaces has enabled us to identify features which affect the relaxation efficiency to the global minimum by considering Landau entropies and free energies. For a surface with a single funnel relaxation is fastest when the potential energy gradient towards the bottom of the funnel is large and the barrier heights are small.

Our simplest model surface appears to capture many of the most interesting features of the dynamical effects seen in clusters and proteins. Our results are also consistent with previous work framed within the language of spin-glass theory; for example increasing the barrier heights raises the glass temperature, T_g, without changing the folding temperature, T_f, and hence decreases the efficiency of relaxation to the global minimum.

We have also modelled kinetic bottlenecks and surfaces with multiple funnels. Multiple funnels can greatly reduce the relaxation efficiency and produce dynamics with both fast and slow time scales. In contrast the effects of kinetic bottlenecks appear to be more limited.

Our interest in this problem was first stirred by work which stressed the importance of the energy gap between the lowest and the next-lowest energy minima on the folding efficiency of model proteins. There immediately appeared to be an obvious analogy with 'magic number' clusters such as the Mackay icosahedron, which are found in simulations and molecular beam experiments. Indeed, one could push the analogy further and consider magic number clusters as those which have survived a selection procedure in a molecular beam, just as proteins have apparently evolved to possess efficient 'funneling' PES's. However, it now appears that factors other than the lowest energy gap are more important. Nevertheless, the analogies between 'magic number' clusters in molecular beams and rapidly folding proteins are probably still valid and will continue to lead to new insights in the global analysis of PES's in terms of the interplay between structure, function, dynamics and thermodynamics.