From Levinthal to pathways to funnels

Ken A. Dill and Hue Sun Chan

A new view of protein folding kinetics replaces the idea of 'folding pathways' with the broader notions of energy landscapes and folding funnels. New experiments are needed to explore them.

A so-called 'New View'[1,2] of the kinetics of protein folding has emerged during the past few years, resulting from a combination of advances in experimental methods that are more informative at the atomic level, and from a new conceptual framework due to simplified statistical mechanical models. In the new view, folding is seen as a parallel flow process of an ensemble of chain molecules. In the metaphor of energy landscapes, folding is seen as more like the trickle of water down mountainsides of complex shapes, and less like the flow through a single gulley. In a nutshell, the new view puts more emphasis on ensembles and multiple folding routes and less emphasis on specific structures and pathways.

The 'New View': what's new?

The groundwork for protein folding kinetics was laid 35 years ago. The famous experiments of Christian B. Anfinsen and co-workers beginning in the early 1960s[3,4] showed that proteins can fold reversibly, implying that the native structures of some small globular proteins are thermodynamically stable states, and therefore are conformations at the global minima of their accessible free energies. But Cyrus Levinthal made the argument in the late 1960s[5–7], later called 'Levinthal's paradox'[8,9], that there are too many possible conformations for proteins to find the 'needle' (the native structure) in the 'haystack' (conformational space) by random searching. Levinthal concluded that proteins must fold by specific 'folding pathways'.

Levinthal framed the puzzle as if the two goals—achieving the global minimum and doing so quickly—were mutually exclusive. These two mutually exclusive options came to be called thermodynamic control and kinetic control. Thermodynamic control meant that a protein reaches its global minimum in energy and that folding is pathway independent (that is, the native structure is determined only by the final native conditions and not by the initial denaturing conditions) but it takes a long time because it requires an extensive search. Kinetic control meant that folding happens quickly (on biological time scales) because it is pathway dependent (that is, the final structure could differ depending on the denaturing conditions from which folding was initiated) and therefore the protein may reach only local optima. Levinthal's argument led to a search for folding pathways.

The view arose that if we could observe intermediate states along the folding pathway, we would learn how nature seives so quickly through the conformational haystack to find the needle—kinetics experiments might teach us the folding code. The modern era of folding experiments began soon after Levinthal's pivotal argument, with the pioneering papers of Ikai and Tanford[10] and Tsong, Baldwin and Elson[11] in 1971 that began the search for folding intermediates. Immediately another puzzle arose: are folding intermediates on-pathway (Tsong *et al.*[11]) or off-pathway (Ikai and Tanford[10])? The implication was that only on-pathway intermediates could teach us the folding code; off-pathway intermediates were uninteresting dead ends. Those papers seeded the large modern enterprise of protein folding kinetics experiments, including the use of disulphide bonding experiments to trap folding intermediates, for example in BPTI (bovine pancreatic trypsin inhibitor)[12–15], and studies of the slowing of folding by proline isomerization[16–19]. But it has become clear that the main story of protein folding kinetics is not told in the details of disulphide bonding or proline isomerization. Protein folding intermediates and complex kinetics can be observed in the absence of disulphide bonds and incorrect proline isomers[20,21].

To appreciate the differences between what Baldwin refers to as the new and classical views[2], we must distinguish carefully between data, on the one hand, and models for interpreting the data, on the other. The new view is based on new models (discussed below): the classical view is based on simple phenomenological kinetics models. What is the basis for the classical models? The raw data are single- or multiple-exponential time decays of optical properties that monitor changes in the protein structure after a jump to folding or unfolding conditions. When a single exponential decay is observed in both folding and unfolding directions, kinetics is called 'two-state' because the relaxation times and amplitudes can be fit by assuming only a native state N and a denatured state D. When multiple exponentials are observed, more complex models are required, with more than two states, so additional symbols are added into mass-action kinetics equations. Each such symbol can stand for an additional

Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-1204, USA

Correspondence should be addressed to K.A.D.
dill@maxwell.ucsf.edu

intermediate conformation (assumed to be distinguishable from native) unfolded, and other intermediate states. Three of the most important classical models are: the off-pathway model, described by the reaction scheme;

$$\begin{array}{l} U \rightleftharpoons N \\ \uparrow\downarrow \\ X \end{array} \qquad (1)$$

the on-pathway model;

$$U \rightleftharpoons X \rightleftharpoons N \qquad (2)$$

and the sequential model;

$$U \rightleftharpoons I_1 \rightleftharpoons I_2 \rightleftharpoons \ldots \rightleftharpoons N, \qquad (3)$$

where U represents the fully unfolded denatured state, N the native state, and X or I represent intermediate states that have properties between those of U and N. Different models imply different interrelationships between relaxation times and amplitudes. Models are chosen based on which one gives the best fit to the experimental rates and amplitudes. The classical experiments generally probe only the average behaviour of the protein, and are not able to resolve much atomic detail.

Although the term pathway appears in both the Levinthal argument about microscopic degrees of freedom and in these macroscopic phenomenological models, this term embodies two very different concepts. Some confusion has arisen from using the single term pathway for both microscopic and macroscopic ideas.

The new view derives from advances in both experiment and theory. The main experimental advances have been those that give faster or more detailed structural information, down to the atomic level. These methods include high-resolution hydrogen exchange (HX)[22–28], particularly pulsed HX[29–31], mass spectrometry[32–34], mutational studies[35–38], and fast laser-triggered methods that explore very early events in folding[39–41]. These new techniques unmask a rich array of information about folding kinetics that was not previously available[42]. An important class of new studies—of small well-characterized fast-folding two-state proteins that do not accumulate any detectable kinetic intermediates[43–48]—are showing how proteins can fold quickly without the guidance of 'midwife' intermediate conformations.

The theoretical advances contributing to the new view are due to a class of models that are fairly new to biochemistry. These are statistical mechanics models with highly simplified—most often lattice-based—representations of chain geometries and interactions, analysed by analytical methods and computer simulations. Although such models lack atomic detail, they go beyond the phenomenological models in including the main microscopic ingredients of proteins: chain connectivity, flexibility, excluded volume, and sequence-dependent intrachain interactions[49,50]. Modelling is not yet feasible for protein folding kinetics at a more atomic level of detail because of computational limitations. Whereas phenomenological models use a single symbol each to represent the denatured state (D), transition state (T) or intermediate state (I), the statistical mechanical models recognize that such macroscopic 'states' are really distributions, or *ensembles* of individual chain conformations. In this view, two macroscopic states 'A' and 'B' are not necessarily distinct sets of microscopic conformations: many of the same molecular configurations may contribute to both macroscopic states. Indeed, such distributions can overlap quite substantially.

The new view replaces the pathway concept of sequential events with the funnel concept of parallel events. The new perspective sees folding as a diffusion-like process[51–54], where the motions of individual chains are asynchronous, each being buffeted by Brownian forces through different sequences of chain conformations, which ultimately all find their ways to the same native structure, in the same way that water flowing along different routes down mountainsides can ultimately reach the same lake at the bottom. But while the new models are described in a language of the microscopic details of the individual chain trajectories (like the trails of small water streams down the hillsides), experiments see global averages over the details (like a photo from an airplane). In common with the classical description, theorists are developing ways to express model folding results as stages of a journey by recognizing the identifiable commonalities among the folding trajectories of the individual molecules. In the mountainside metaphor, modellers are finding ways to recognize when water streams come together to run through common gulleys, for example. Some hydrophobic clustering may occur early, one secondary structure may form before another, a disulphide bond may happen later, and so on. It is this description of protein folding as stages of a journey that comes closest to capturing experimentalists' meaning of the the term pathway.

The new models use the language of 'folding funnels' and 'energy landscapes'[51–65]. An energy landscape is just the free energy of each conformation as a function of the degrees of freedom, such as the dihedral bond angles along the peptide backbone. The vertical axis of a funnel (Figs 1–6) represents the 'internal free energy' of a given chain configuration (that is, everything except the conformational entropy): the sum of hydrogen bonds, ion-pairs, torsion angle energies, hydrophobic and solvation free energies, and so on, for a chain in a particular conformation. (Because such terms depend on external conditions, the vertical extent of the energy surface will stretch or shrink with temperature or changing solvent). The many lateral axes represent the conformational coordinates. The high dimensionality of this representation reflects the many degrees of freedom of a protein chain. Many coordinates, for example the dihedral angles ϕ_1, ψ_1, ϕ_2, ψ_2, ..., are needed to specify a conformation (only two are shown schematically in Figs 1–6). Each conformation is represented by a point on the multidimensional energy surface. Conformations that are similar geometrically are close to one another on the energy landscape. Energy landscapes can have many kinds of features. Hills correspond to high energy conformations (for example burying polar groups in hydrophobic cores, or unfavourable $\phi\psi$ angles), and valleys are configurations that are more favourable than others nearby.

The kinetic process of folding or unfolding a protein can be likened to rolling a ball on this energy surface. Upon initiation of folding conditions, the protein tends to change its conformation in ways that cause its energy to decrease, but it is also constantly buffeted into different conformations by Brownian motion. In this analogy, each individual protein molecule corresponds to a ball rolling on the energy landscape following some particular trajectory, winding through the hills and valleys, all the while being randomly redirected by Brownian forces. Uphill steps happen too, but less often.

The landscape picture suggests caution in the use of terminology like pathway, on-pathway and off-pathway, kinetic and equilibrium control, intermediate state, and transition state, because such terminology evokes thinking about specific structures rather than ensembles, and this sometimes (but not always) leads to conceptual quagmires, as noted below.

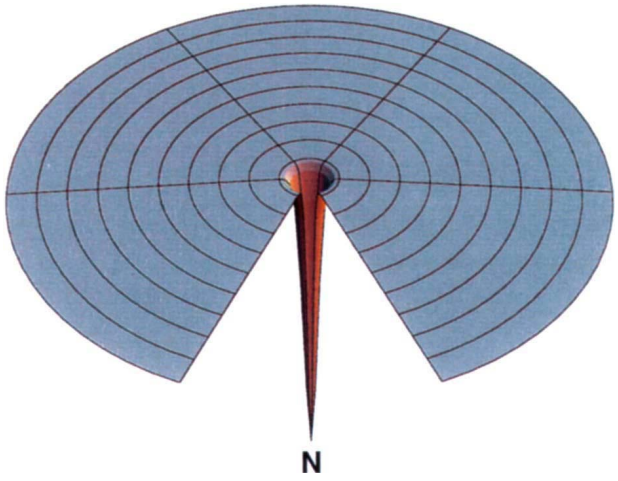

Fig. 1 The Levinthal 'golf-course' landscape. **N** is the native conformation. The chain searches for **N** randomly, that is, on a level playing field of energies.

For example, consider the idea of a 'folding pathway'. Figs 1–3 illustrate the conceptual difference between pathways and funnels, using the energy landscape metaphor. Levinthal's argument that random searching would not find the native states corresponds to the hypothetical 'flat playing field' or 'golf-course potential' shown in Fig. 1. When a ball rolls randomly on a flat course, it takes a long time to find, and fall in, the hole. From this perspective, as Levinthal notes, proteins would have a serious search problem.

Fig. 2 uses the energy landscape metaphor again to show how Levinthal envisioned that pathways could solve the search problem of Fig. 1. Beginning from a denatured conformation **A**, a pathway embodies the idea that the folding molecule goes through a sort of tunnel on the landscape, like water flowing down a gutter, to the native structure **N**. This process is more directed than a random search. According to this idea, 'a pathway of folding means that there exists a well-defined sequence of events which follow one another'[5]. The gutter represents a particular series of changes in dihedral angles. It may have valleys (intermediate states) and/or hills (transition states) on its way to the native state.

The concepts of on-pathway and off-pathway intermediates have their roots in images like Fig. 2, which are defined by whether such valleys are contained within the gutter, or outside it, respectively. But while the pathway idea shown in Fig. 2 handily 'solves' the search problem embodied in Fig. 1, the physical basis for such specific sequences of events is unclear. Moreover, this picture creates an artificial problem, namely the Levinthal dichotomy of thermodynamic *versus* kinetic control, pathway dependence *versus* pathway-independence. The new view recognizes that the fundamental problem with Levinthal's solution is the concept of 'pathway' itself.

The Levinthal paradox is not a real problem. The 'paradox' is little more than a misconception about how any physical, chemical, or biological system that is governed by thermodynamics can reach its stable states in measurable times. Thermodynamics texts are full of examples of systems having nearly Avogadro's number of microscopic degrees of freedom that nevertheless reach stable states on observable time scales. The two goals of reaching a global energy minimum and doing so quickly are not mutually exclusive. The paradox is an artifact of framing the folding problem in terms of the landscape of Fig. 1. A pathway can lead from a specific point **A** to a specific point **N**, as in Fig. 2. But folding a protein does not involve starting from one specific conformation, **A**. The denatured state of a protein is not a single point on the landscape: it is all the points on the landscape, except for **N**. A pathway is too limited an idea to explain the flow from everywhere else, the denatured ensemble, to one point **N**. The concept of a pathway is useful for explaining the milestones we see in travels along a road or along a hiking trail, but not for describing how rain flows down a funnel.

The new view recognizes that the solution to Levinthal's paradox is 'funnels, not tunnels'[55]. This view has arisen from work on several different models. Folding landscapes do not look like Fig. 1. Fig. 3 shows an idealized smooth protein folding funnel based on an early mean-field lattice model[8]. Bryngelson and Wolynes first explored the bumpiness of protein folding landscapes in simplified spin-glass based models[56,66]. Leopold, Montal and Onuchic[67] first described in some detail how the shape of a folding funnel depends on amino acid sequence, by computer enumeration of conformations in lattice heteropolymer models. Since the lateral area of an energy landscape at a given depth represents the number of conformations (or conformational entropy) having the given intrachain free energy, the funnel idea is simply that as a folding chain progresses toward lower intrachain free energies—by increasing compactness, hydrophobic core development, intrachain hydrogen bonding, salt-bridge formation, and so forth—the chain's conformational options become increasingly narrowed, ultimately toward the one native structure. Fig. 3 is an idealization (a landscape without features or bumps) showing how the many open conformations might funnel down through the fewer compact conformations, and finally to the one native conformation.

Fig. 3 shows how funnels resolve Levinthal's paradox. We can draw an analogy between the denatured 'state' and an ensemble of skiers distributed over a mountainside. When folding conditions are initiated, each skier proceeds down the funnel following his own private trajectory. Skiers skiing down funnels reach the global minimum (satisfying Anfinsen's thermodynamic hypothesis), by many different routes (not a single microscopic pathway), yet

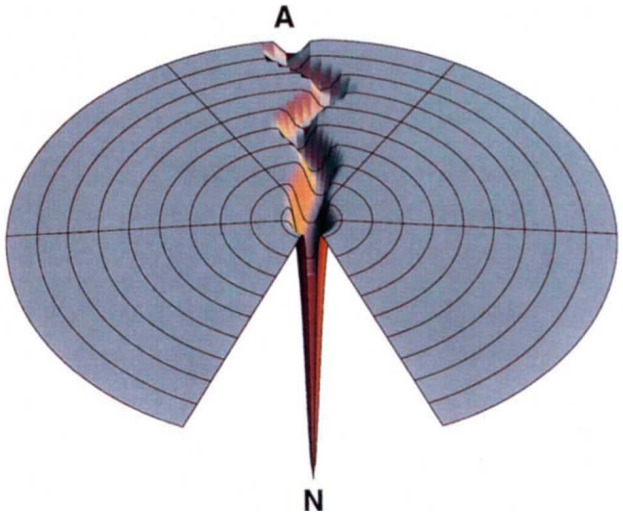

Fig. 2 The 'pathway' solution to the random search problem of Fig. 1. A pathway is assumed to lead from a denatured conformation **A** to the native conformation **N**, so conformational searching is more directed and folding is faster than for random searching.

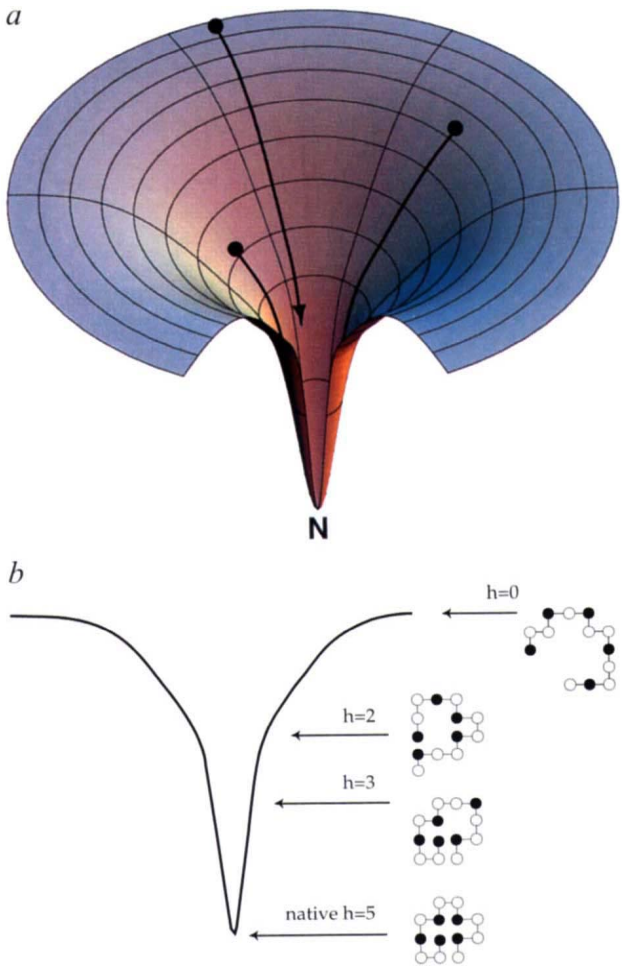

Fig. 3 *a*, An idealized funnel landscape. As the chain forms increasing numbers of intrachain contacts, and lowers its internal free energy, its conformational freedom is also reduced. *b*, A slice through (*a*). In the lattice model, black beads represent hydrophobic monomers and white beads represent polar monomers, *h* is the number of hydrophobic contacts. Exact enumeration studies show that there are many open conformations, fewer compact conformations, and only one conformation having *h*=5. An ensemble of molecules can reach the global minimum of free energy (satisfying Anfinsen's thermodynamic hypothesis), and do so quickly (satisfying Levinthal's concern), even though each chain follows its own route, not a single pathway.

they do so in a directed and rapid way (satisfying Levinthal's concern). Each different denatured conformation progressively reconfigures rapidly toward the native conformation, even though individual molecules can each follow different sequences of dihedral angle changes. By eliminating pathway terminology, this picture eliminates artifactual paradoxes about how folding can be both 'pathway'-dependent and 'pathway'-independent at the same time. While pathways lead us to think in terms of the sequential assembly of specific structures, funnels lead us to think in terms of parallel processes and ensembles.

Folding funnels may not be generally as smooth and simple as shown in Fig. 3. Fig. 4 indicates an artist's rendition of another possible scenario. The distinction between a pathway and a funnel is that a pathway is a one-dimensional route through configuration space, whereas a funnel describes the progressive reduction in dimensionality of the accessible conformational space, beginning from the many degrees of freedom available to denatured chains, ultimately down to the nearly complete lack of freedom of the native chain. In short, the funnel metaphor refers not only to shapes of real funnels; it refers more broadly to the overall narrowing of conformational options in the downward direction as the internal free energy of a protein chain decreases. Whereas the pathways of classical models are hiking trails, the new view proposes that multi-dimensional energy landscapes of proteins can have a much broader array of shapes, as mountain ranges do, that involve hills, valleys, ridges, channels, moguls, plains, valleys inside valleys, moats, varying slopes, and ups and downs of all kinds, and higher-dimensional features that are more difficult to describe. Perhaps a more accurate description of what's new in the new view is to refer to it as the Landscape Perspective and the earlier view as the Pathway Perspective. In this sense, what is new is really just a broadening of the possibilities for how proteins might fold.

The new view of the transition state

Two-state fast folding kinetics is described by funnel-shaped landscapes with no significant kinetic traps. Slow multi-exponential folding is represented by bumpy or rugged landscapes, like the Himalayan mountains, where it is very difficult for skiers to reach the deepest valley directly, because they tend to get caught in bottoms of valleys at altitudes higher than the lowest point on the mountain range.

In classical chemical kinetics, 'transition state' is a term that describes a rate-limiting step. Applied to small-molecule reactions, it is a high-energy configuration that is challenging for a system to achieve. It is the top of the hill along the reaction pathway, the highest point of a mountain pass on the potential energy surface. The transition-state picture from small molecule chemistry has provided a historical backdrop and a structural perspective for interpreting protein folding kinetics. From that perspective, we might expect the transition state of folding to correspond to some particular 'bottleneck conformation', a hill on a specific reaction pathway. But while the new view replaces pathways with mountainsides of all types, it also replaces transition states as highest bumps on a trail with a broader view of the transition states as whatever are the bottleneck processes of flows down different mountainsides. The new perspective recognizes that the kinetic bottleneck for folding does not necessarily describe specific contortions or particular structures of the chain, although the new view does not preclude them[51]. In the new view, 'transition state' is fundamentally a concept about rates, not about specific structures. If rain flows down a single gulley on a hillside, the bottleneck may be a particular obstruction in the gulley. But if rain washes down a complex mountainside, then defining its 'reaction coordinate' is difficult, and there may be no single bottleneck. The overall flow rate may be the integral over many smaller processes.

Bumpy landscapes lead to multi-state kinetics

What are the bottlenecks to folding? We return to the analogy of skiers on a mountain. Under denaturing conditions, an ensemble of skiers is distributed everywhere around the funnel mountainside. A stopped-flow jump to folding conditions steepens the hillsides and causes all the skiers to proceed downhill. Each may follow a different route. After skiing down a mountainside like Fig. 4, skiers enter a rugged landscape of hills and valleys and kinetic traps at the bottom, corresponding to the accumulation of 'misfolded' intermediates. For these proteins, the slow steps arise from climbing an uphill slope (breaking existing favourable contacts) after being trapped in local minima (transient intermediates), then reaching a mountain pass, before returning to the next downhill search[51–54,58–61,63,68–73] (Figs 4, 5). Folding may proceed in two or more kinetic phases, often with a fast collapse to a com-

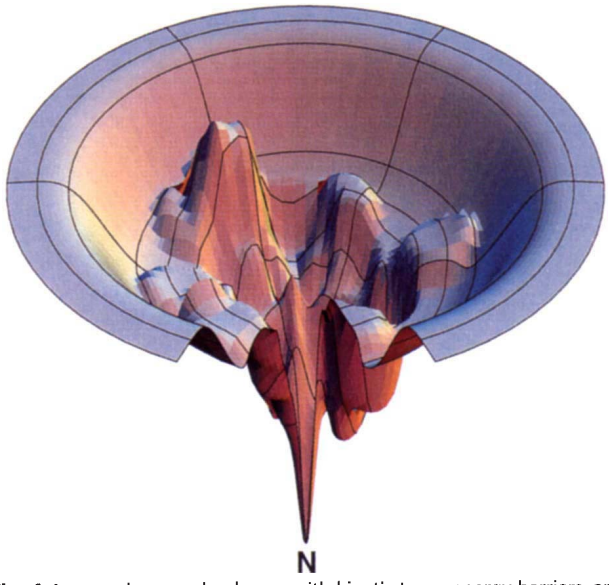

Fig. 4 A rugged energy landscape with kinetic traps, energy barriers, and some narrow throughway paths to native. Folding can be multi-state.

pact ensemble (skiing down to the traps) followed by slow reconfiguration of kinetically trapped compact non-native conformations into the native structure (uphill climbing to 'lookout points' and skiing down again). In these cases, the transition 'state' is the ensemble of lookout-point conformations that have been opened up and pulled apart, relative to the compact trapped states from which they originated[58,59,74].

Hence the folding transition state can be many different chain conformations. No single mountain pass or lookout point in the Himalayas can be considered to be the bottleneck. From a given valley, a chain can break different contacts (climb different hills) increasing its conformational entropy, before proceeding downhill again. This does not necessarily imply that folding barriers are enthalpy-controlled in these cases, because the total experimental entropy includes other contributions than just chain conformational entropy. Uphill climbing does not necessarily mean a full opening of the chain; sometimes just a few contacts here or there need to be broken for the chain to resume progress toward the native state.

The landscape perspective sees intermediates somewhat differently than is implied by the terms 'on-' or 'off-pathway'. Defining intermediates as on-pathways *versus* off-pathway has often been done based on energies, rather than on structures. That is, intermediates are called off-pathway if such conformations must break bonds, go up hill, and thus become more unfolded in an energetic sense before proceeding toward the native state. But on-pathway *versus* off-pathway could also be defined in structural rather than energetic terms. For example, when β-lactoglobulin, a predominantly β-sheet protein passes through a helical state as it folds[75,76], this is off-pathway in the structural sense that the folding ensemble does not monotonically increase its resemblace to the native structure.

Energy landscapes show that the question of whether a route is as direct as possible on a landscape is different from whether it involves uphills or downhills. For example, Fig. 5 shows a Moat Landscape, indicating a funnel-like 'throughway' path for the A routes and an obligatory kinetic trap for the B routes[59], a splitting that has been referred to as 'kinetic partitioning' by Thirumalai[70]. This may correspond to hen egg white lysozyme, studied by Chris Dobson and his colleagues[31,33], in which there is a subpopulation of chains that undergo overall fast folding (the A routes), and another subpopulation that forms the α-helical domain quickly but the β-sheet domain slowly (the B routes). In classical terminology, the Moat intermediate shown in Fig. 5 would be called off-pathway, since the chain exits the moat trap by an uphill step and breakage of contacts, partially denaturing before refolding. But in the landscape view, these traps are as direct and 'on-route' as is possibly achievable for that part of the chain population. Moreover, in the landscape picture, intermediates are 'slowing-down places' on mountainsides—the chain gets stuck in moats, behind hills, lost in meadows, trapped in moguls or *cul-de-sacs*, and so forth, a much broader spectrum of options than is implied by the binary choice between on-pathway and off-pathway.

Another conceptual puzzle that is readily rationalized within the new view regards chaperonin proteins. From a structural perspective, the central question of Hsp 60 chaperonin action might be: how can a single type of chaperonin protein, such as GroEL, 'recognize' the transition state structures of different substrate proteins—rubisco, dihydrofolate reductase[77,78], and others? From the classical perspective that emphasizes specific structures, one expects the folding transition states to be different from one substrate protein to the next. But the landscape perspective shows how folding can be accelerated without 'molecular recognition'. To catalyse folding, a chain may merely need to be pulled apart nonspecifically, bringing it uphill on an energy landscape[79–82]. One way to help a skier find the lowest point on a mountain range is to just keep dragging him uphill, in random directions by trial and error, and let him ski down again. This strategy is universal for any protein. It doesn't depend on the shape of a mountain range or an energy landscape. That is, pulling apart one protein to let it attempt to refold is not much different than pulling apart another, a process that has been called 'iterative annealing'[80,81].

Some smoother funnels describe two-state kinetics

For fast two-state kinetics, involving no significant kinetic traps, the bumps and ruggedness are probably much less important than for multistate folding. For these cases, the mountainside is more like a funnel, flat at the top and steeper and deeper toward

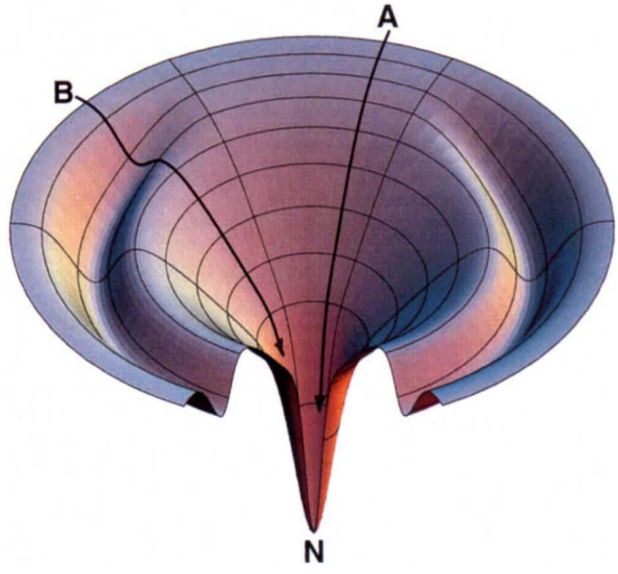

Fig. 5 Moat Landscape, to illustrate how a protein could have a fast-folding throughway process (A), in parallel with a slow-folding process (B) involving a kinetic trap.

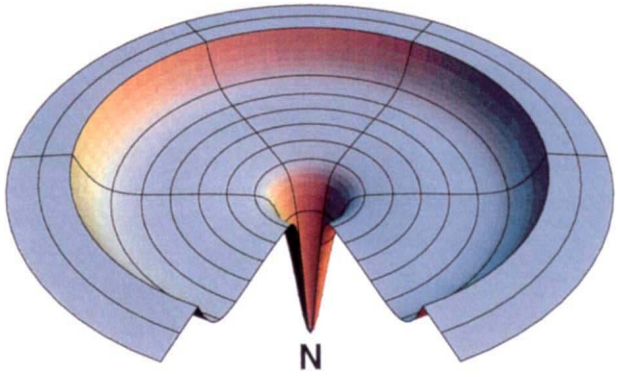

Fig. 6 Champagne Glass Landscape, to illustrate how conformational entropy can cause free energy barriers to folding. The 'bottleneck' or rate limit to folding is the aimless wandering on the flat plateau as the chain tries to find its way downhill.

the native structure. What is the rate-limiting process for such two-state folding? What slows down each skier is just the snow along the way. Skiing to the bottom of the hill is intrinsically limited by the size and slope of the hill, the number of obstacles, and the Brownian buffeting. Just as skiing down any mountain takes time, even when there is no specific obstacle that can be held solely responsible, likewise fast two-state folding can be a process in which the chains slosh around in solvent seeking better structures, with no specific bottleneck conformation. A chain configuration that is called 'denatured' at the beginning of the process (some point on the landscape) can also be visited by other chains as they flow down to the native state.

Since the idea of 'transition state' is really about rate limits and bottlenecks, it includes all the conformations that are passed through on the way to the native state[73], because they are all responsible for determining the rate. Thus the transition state ensemble and the denatured state ensemble are identical in this idealized case! This does not preclude the possibility that different conformations may have quite different effects on the overall folding rate. It just means that the rate limiting steps are not necessarily confined to the formation of a highly restricted set of conformations. Model calculations show that folding on some smooth funnel landscapes is two-state: collapse and folding essentially coincide in an approximately single-exponential process[69,73,83–86]. Chain organization can be quite a complex function of depth in landscapes, even in simple model landscapes. Kinetics on model smooth funnels can sometimes be more complicated (N.D. Socci, personal communication).

For two-state kinetics, transition states may involve free energy barriers due to conformational entropy. What are conformational entropy barriers? Fig. 6 shows a Champagne Glass Landscape, to illustrate bottlenecks due to conformational entropy. A skier is delayed en route to the bottom by aimless searching on the flat meadow to find the remaining route downhill. That is, the chain is lost in configuration space and must spend time searching for ways downhill. (This corresponds to a process that goes 'uphill in free energy' to overcome a 'free energy barrier' in conventional terminology, which just means that the process is slow relative to some reference rate.) For protein folding, this might correspond to the conformational meandering of a long loop of polar residues that will ultimately bring two hydrophobic clusters together. The idea of entropy barriers is not new: diffusion-controlled processes have been studied for many years[87–90].

How do energy landscapes help us interpret experiments? The central experimental results on two-state kinetics are chevron plots of denaturant effects on rates, and Arrhenius plots of temperature effects on rates. In the energy landscape picture, changing the solvent or temperature toward more native-like conditions has the effect of stretching down the energy surface as if it were a rubber sheet.

Under weakly folding conditions, the funnel is shallow (the native structure is not strongly favoured by intrachain interactions), so at equilibrium the protein tends to spend most of its time meandering around the upper part of the funnel where there are more accessible chain conformations, and folding is slow because the chain meanders quite aimlessly down the shallow slope. But under more native-like conditions, the funnel is stretched down. As a result, the Brownian buffeting is not strong enough to keep the protein in the upper part of the funnel for long, and most of the molecules move down towards the native state. Folding is fast because the slope is steep, and skiers are directed strongly toward the bottom.

Fig. 7 shows predictions of chevron plots and temperature-dependences from lattice simulations, and corresponding experiments[35,91–94] for qualitative comparison. The statistical mechanical models can rationalize non-Arrhenius rate laws[51,59,60,63,73], chevron plots[73], mutational effects[51,59,63,95], kinetic traps, barriers[51,58–61,63], chaperonin action[79–82], the apparent 'pathways' in some protein folding experiments[51], and the relation of equilibrium properties and fluctuations to kinetics[51–54,61,62,71,85,86,96–98].

How is the shape of an energy landscape determined by the amino acid sequence? One hypothesis is that collapse happens by 'hydrophobic zipping'[63,99], an opportunistic process in which local contacts (those nearby in the sequence) form first, drawing in new contacts, which create still other and increasingly nonlocal contacts, and opportunities for other intrachain interactions, and so on. Helices and turns and other local structures would be the first to zip. As the chain moves ever lower in internal free energy, the developing intrachain contacts reduce the conformational entropy. Hydrophobic zipping is nothing but steepest descent on landscapes, because it defines routes involving minimal entropy loss per step downhill on the energy landscape.

Do proteins fold by nucleation? Theoretical modelling is not yet clear on this point. Recently, it has been proposed that folding of some small proteins such as chymotrypsin inhibitor 2 (CI2) proceeds by a nucleation mechanism, with a highly specific nucleus as the transition state[37,38,44,100]. However, a more recent study of mutational data and model calculations shows that the kinetic bottleneck in CI2 folding may be better described as a broad ensemble of conformations in which different native contacts have different degrees of participation[101].

What is notable about the transition states of folding, according to the new view, is not that they are specific structures, but that they are ensembles[58,59,62,73]. The classical view focuses on specific structure (which experiments see), whereas the new view is an ensemble perspective that recognizes the importance of disorder and that random processes and wrong steps are also major contributors to folding speed. According to the new models, small changes in the solvent or temperature can often lead to continuous subtle changes in the populations of these molecular configurations. Such subtlety is not recognized in the classical modelling, where whole new symbols like I_1 or I_2 are invoked whenever data is not accounted for by simpler models.

The new models

We all use models—not just theorists, but experimentalists as well. When we write down an equilibrium constant K for any reaction $A \rightleftharpoons B$, or a mass-action rate law, or when we postulate that K is a ratio of rate constants, we are using models and making assumptions that are often implicit and seemingly innocuous. Even simple equilibrium and mass action expressions are not direct transcriptions of data. They are full of implications, and based on assumptions: sometimes about two states, or assumed intermediates, or about pathways. Or they neglect certain complications, like side reactions or nonidealities, or they suppose denatured states are open random confirmations, or they assume the constancy of equilibrium and rate 'constants', or they approximate or neglect the temperature dependences of enthalpies or heat capacities, or they assume idealized limiting laws (ideal gas law, Raoult's law, Henry's law, and so on)—even when reality is more complex.

The new view is based on models that do not assume single or multiple exponential behaviour. These models do not assume distinct identifiable macroscopic states, to which a label such as I_1, I_2, U, D and so on, can be attached. The new models do not assume that macroscopic states are independent of conditions (that is, unchanged by temperature, denaturants, pH, salt, and so forth). The new models do not assume pathways or intermediates, on- or off-pathway.

The new models differ from the classical models in capturing some of the molecular nature of proteins. The new models recognize the polymeric connectivity of the chain; that proteins are made of different monomers and have sequence-specific interactions; that protein compactness is limited by excluded volume; that the native states of proteins are often unique stable states. They permit unbiased explorations of the full ensemble of all conformations available to the chain. Even though lattice models treat larger numbers of entities—the many chain conformations—than do phenomenological models, which treat only a few macroscopic-symbol states, the statistical mechanical models tend to use fewer adjustable parameters. Phenomenological multi-exponential models tend to have at least six adjustable rate and amplitude parameters. Some minimal lattice models use only one parameter[49,63].

Of course the new models also have limitations, assumptions and simplifications: they neglect atomic detail, they involve shortened chains, simplified energies and chain representations, and sometimes they are two-dimensional (Fig. 7). Reaching beyond the new view will require a next generation of models that can still broadly explore conformational and sequence spaces while being more faithful to structural details.

Modelling using landscape ideas is only at a very elementary stage. The perfect funnel is an idealization, like the ideal gas law—probably not truly applicable to any real protein, but useful as a 'limiting-law' tool for conceptualizing chevron plots and other aspects of folding. Figs 1–6 are artist's renditions, not results from any actual simulation. The landscape metaphor is mainly useful for illustrating principles, designing new experiments, and posing new questions, not for details. No theory or simulation yet computes the folding kinetics for lysozyme. Modelling does not yet show whether the rate limiting step in folding is the same as in unfolding[93,102,103]. There is not yet agreement on how to establish a single reaction coordinate. It is not fully clear how mutations might affect the apparent 'position' of the transition state along the reaction coordinate[37,38,44]. Nor is it clear whether nucleation is a property of a particular model or a particular amino acid sequence. Much remains to be learned about how to relate kinetic to thermodynamic intermediates. It is not really clear what constraints there are on our use of 'artist's renditions' when we draw fictitious landscape pictures. The advantage of the classical view was the small number of models to choose from. The disadvantage of the new view is the virtually unlimited array of energy landscape shapes and physical models that underlie them. The

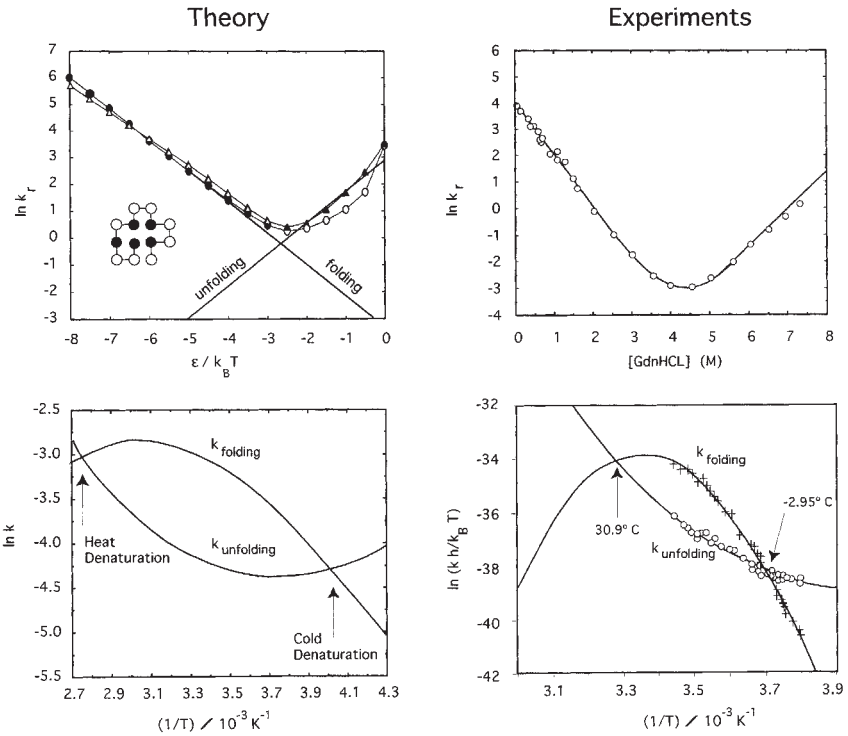

Fig. 7 The chevron plot (upper) and temperature (T) dependence of folding and unfolding rates (lower), calculated from a two-dimensional square-lattice model with the sequence shown in its unique native structure are shown on the left. The chevron plot (upper) and temperature dependence of folding and unfolding rates from experiments[35,43] (lower) are on the right. In the model, there is no one particular structure that is the transition state barrier. In the 'HP+' model used in the panels on the left, every native hydrophobic-hydrophobic contact is favoured by ε (<0), other native contacts are neutral, and every nonnative contact is disfavoured by −ε (details of this model are given in ref. 73). k_r is relaxation rate, h is Planck's constant, and k_B is Boltzmann's constant. Circles and triangles in the upper left panel represent folding and unfolding, respectively. Filled symbols are folding and unfolding relaxation rates measured under conditions that are favourable for folding and unfolding, respectively. The extrapolated straight lines from the chevron plot show approximate folding and unfolding rates as a function of hydrophobic contact energy. Since the hydrophobic contact energy varies approximately linearly with concentration of the denaturant GdnHCl (refs. 49, 106) the model results in the upper left may be compared to experimental chevron plots of relaxation rates versus denaturant concentration. The upper right panel shows the chevron plot for chymotrypsin inhibitor 2 at 25 °C (adapted from ref. 43). In the lower left panel, temperature dependence of the model folding and unfolding rates (k) are computed using the temperature dependence of the hydrophobic interaction[107]. The lower right panel shows experimental results for T4 lysozyme (adapted from ref. 35).

enterprise of model building has only begun to scratch the surface, for determining how the landscape shape (and thus the folding dynamics) depends on physical interactions, model, and amino acid sequence. While these models address some principles of folding kinetics, they are not yet recipes for predicting native structures from amino acid sequences.

A wish list for experimentalists

How can we test the new view, or disprove it and move on to the Even Newer View? The new view is not a single model or prediction; it is a perspective and a language based on energy landscapes. Here we propose a wish list of experiments that could help determine the shapes of protein folding landscapes.

1. Distinguishing pathways from funnels. At the top of our wish list are experiments that could describe the correlations among the many degrees of freedom within a protein molecule. For illustration, consider a property θ_i, for amino acid i=1, 2,...., n in a protein having n amino acids. For example θ_i might be a backbone dihedral angle, ϕ_i or ψ_i. We could distinguish pathways from funnels by measurements of correlations among such properties. That is, if a protein folds by a pathway, then whenever an amino acid, say number 10 in the sequence, has a particular value of its bond angle $\theta_{10}= \alpha_0$, then all the other bond angles in the chain will also have particular characteristic values. But if proteins fold by idealized perfect funnels, then the fluctuation of θ_{10} around its mean value at a given time during folding will not be correlated with the corresponding fluctuation of θ_{27} or any other bond angle. In reality, we expect correlations between fluctuations of variables such as θ_i and θ_j in real proteins to be intermediate between these two extreme cases, and to change with time as folding proceeds.

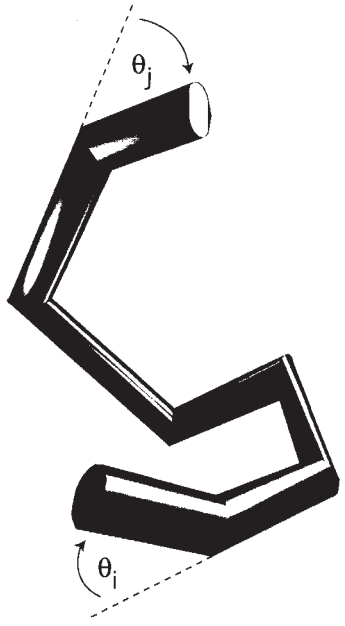

Fig. 8 Measuring correlations, for example between the fluctuations of two bond angles θ_i and θ_j, should help determine whether degrees of freedom are synchronized, as in the pathway model, or asynchronous at first and becoming more synchronized later.

Measuring these correlations requires measuring distributions of populations, not commonly done, not just measuring the averages observed in most experiments. That is, since folding is cooperative, it means that the ensemble average value $\langle\theta_{27}\rangle$ will naturally proceed toward its native value over the same time course that $\langle\theta_{10}\rangle$ proceeds towards its native value. This is not the correlation we seek. Rather, we ask whether the fluctuations in θ_{27} correlate with the fluctuations in θ_{10}, that is, whenever there is a transient increase in θ_{10}, there is, say, a transient decrease in θ_{27}. In mathematical terms, we seek the ensemble averages of correlation quantities such as $\langle\theta_i\,\theta_j\rangle$ (Fig. 8). The aim is to observe how folding reduces the dimensionality of the conformational space, and how the different degrees of freedom become increasingly correlated until no freedom remains. The funnel perspective expects few correlations early in the folding process, and increasing amounts or correlation as folding proceeds.

Our description above applies to some property of each monomer i, like a dihedral angle. Other examples of monomer properties include the solvent environment or degree of burial, which can be measured by fluorescence or hydrogen exchange protection. Of equal interest are pairwise quantities like contacts or distances between monomers i and j. In this case, is a contact, say (i, i+3) correlated with some other contact, say (j, j+3)? That is, when one occurs, does the other also invariably occur? Irrespective of what property is measured, the main point is that the distinction between the new and classical views will be found in some measure of the correlations among chain properties, rather than in their average values alone. A promising prospect is the combination of mass spectrometry experiments in conjunction with hydrogen exchange labelling[32–34].

2. Barriers on the landscapes. While the folding process is a transition from one macroscopic state to another, say denatured to native, we hope experimentalists will develop ways to measure transition rates among more microscopic subpopulations, for example between the native and 'first excited' states. The conformations of proteins can be classified in analogy with quantum mechanical energy ladders, with the native state at the bottom rung. The first excited conformations are slightly higher in internal free energy, due to some broken noncovalent bonds. The importance of the first excited conformations of proteins is that they are the main thermal fluctuations of the protein under native conditions. Second excited state conformations are at the next higher rung on the energy ladder, and so on. NMR measurements of the rates of hydrogen exchange can determine the populations of the first few excited states, based on the EX2 exchange mechanism[22–25,104]. To determine the kinetic barriers on the energy landscapes, we hope to know not just the populations of these excited states, but also the transition rates between them. Perhaps such data could be obtained by the so-called EX1 mechanism[105], or by other methods.

Another experiment on our wish list is a way to probe the bumpiness of energy landscapes. If a landscape is a smooth funnel, then small fluctuations (from native) in energy must lead to small changes in structure; see the left side of Fig. 9. This is the traditional view of protein fluctuations as 'small wiggles'. But a bumpy landscape means that when a native molecule undergoes its normal thermal fluctuations, which are small in energy, it could undergo large changes in structure to "conformational distant relatives"[105]. Perhaps a β-sheet protein occasionally fluctuates into a helix. To search for conformational distant relatives will require new types of experiments. If the first excited conformations of a protein include many small wiggles and a few conformational distant relatives, then the averages over the first excited states such as those seen by HX will be very insensitive to the distant relatives. We hope for ways to trap individual fluctuations, particularly those that are very non-native in conformation, and that are only fleetingly present under native conditions. Another way to map out the sizes and locations of basins on energy landscapes is to explore how the folding kinetics depend on various different initial conditions[45,48].

3. The view from the high vistas. The tops of folding funnels have been difficult to study because the first stages of folding happen faster than the millisecond deadtimes of most stopped-flow experimental measurements. Until recently, most experiments explored mainly the big barriers and the bottoms of funnels. But faster experiments in which laser temperature jumps synchronize folding on nanosecond and microsecond time scales have recently become feasible[39–41]. Unfortunately, current detectors give little

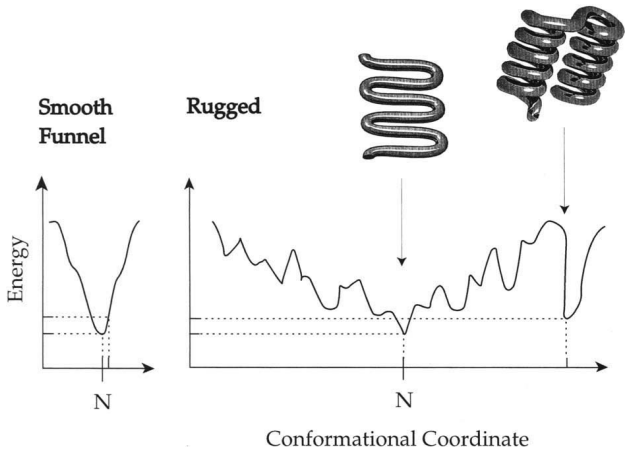

Fig. 9 If landscapes are smooth, then native proteins should have small fluctuations, but if landscapes are rugged, then native proteins could fluctuate to very different conformations. Small changes in energy could lead to large changes in structure.

atomic detail. Another item on our wish list is high-resolution data from these fast experiments.

4. Skiing down the slopes. What are the slopes of landscapes? What limits the rates of the fastest folding proteins? Eyring theory for the reaction rate k of small molecule chemical reactions,

$$k=\frac{k_B T}{h}\exp\left(-\frac{\Delta G^{\ddagger}}{k_B T}\right) \qquad (4)$$

is sometimes applied to protein folding, where k_BT is Boltzmann's constant times absolute temperature and h is Planck's constant. The front factor, k_BT/h, in this expression describes the fastest velocity the reaction could achieve when there is no free energy barrier, that is when $\Delta G^{\ddagger}=0$. For chemical reactions in the gas phase, the limiting rate is the speed of a single bond vibration, hence the factor k_BT/h. But the Eyring theory, and the front factor k_BT/h, are not appropriate for protein folding, which involves a collapsing molecule in a solvent, not a chemical bond formation in a vacuum. Other rate theory models are available for small molecule reactions, that indicate how the front factors might depend on solvent viscosities[87–90]; perhaps generalizations of these will be useful for proteins. We hope experimentalists will tell us the following. What is the fastest speed a protein can fold? How does the fastest folding speed depend on viscosity? This will help discriminate folding processes that happen while the chain is solvated, from those that happen after the chain has excluded solvent. Is the maximum folding speed different for helix bundles versus β-sheet proteins, or other tertiary structures? This should lead to better theoretical models for the rate-limiting slopes of energy landscapes.

Summary

While the classical view of protein folding kinetics relies on phenomenological models, and regards folding intermediates in a structural way, the new view emphasizes the ensemble nature of protein conformations. Although folding has sometimes been regarded as a linear sequence of events, the new view sees folding as parallel microscopic multi-pathway diffusion-like processes. While the classical view invoked pathways to solve the problem of searching for the needle in the haystack, the pathway idea was then seen as conflicting with Anfinsen's experiments showing that folding is pathway-independent (Levinthal's paradox). In contrast, the new view sees no inherent paradox because it eliminates the pathway idea: folding can funnel to a single stable state by multiple routes in conformational space. The general energy landscape picture provides a conceptual framework for understanding both two-state and multi-state folding kinetics. Better tests of these ideas will come when new experiments become available for measuring not just averages of structural observables, but also correlations among their fluctuations. At that point we hope to learn much more about the real shapes of protein folding landscapes.

Note added in proof: Our wish number 4 is already being addressed[108,109]. An upper limit of ~(1μs)$^{-1}$ on the rate of protein folding has recently been proposed based on new experiments on the rate of intrachain diffusion[108].

Acknowledgements

We thank R.L. Baldwin, S. Bromberg, C. Frieden, T.G. Oas, J. Schonbrun, J. Schreurs, N.D. Socci, T.R. Sosnick, and D. Thirumalai for very helpful comments on this paper. We thank S.W. Englander, Y. Goto, C.R. Matthews, J.N. Onuchic, and J.A. Schellman for stimulating discussions, and D.S. Heap and J. Schreurs for creative and technical help with the figures. We thank the National Institutes of Health for financial support.

1. Baldwin, R.L. Matching speed and stability. *Nature* **369**, 183–184 (1994).
2. Baldwin, R.L. The nature of protein folding pathways: The classical versus the new view. *J. Biomolec. NMR* **5**, 103–109 (1995).
3. Anfinsen, C.B., Haber, E., Sela, M. & White, F.H. Jr. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. USA* **47**, 1309–1314 (1961).
4. Anfinsen, C.B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
5. Levinthal, C. Are there pathways for protein folding? *J. Chim. Phys.* **65**, 44–45 (1968).
6. Levinthal, C. in Mossbauer Spectroscopy in Biological Systems. Proceedings of a meeting held at Allerton house, Monticello, Illinois. (eds P. Debrunner, J. Tsibris, & E. Munck) 22–24 (University of Illinois Press, Urbana, Illinois, 1969).
7. Wetlaufer, D.B. Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl. Acad. Sci. USA* **70**, 697–701 (1973).
8. Dill, K.A. Theory for the folding and stability of globular proteins. *Biochemistry* **24**, 1501–1509 (1985).
9. Zwanzig, R., Szabo, A. & Bagchi, B. Levinthal's paradox. *Proc. Natl. Acad. Sci. USA* **89**, 20–22 (1992).
10. Ikai, A. & Tanford, C. Kinetic evidence for incorrectly folded intermediate states in the refolding of denatured proteins. *Nature* **230**, 100–102 (1971).
11. Tsong, T.Y., Baldwin, R.L. & Elson, E.L. The sequential unfolding of ribonucleases A: Detection of a fast initial phase in the kinetics of unfolding. *Proc. Natl. Acad. Sci. USA* **68**, 2712–2715 (1971).
12. Creighton, T.E. Conformational restrictions on the pathway of folding and unfolding of the pancreatic trypsin inhibitor. *J. Mol. Biol.* **113**, 275–293 (1977).
13. Creighton, T.E. Experimental studies of protein folding and unfolding. *Prog. Biophys. Mol. Biol.* **33**, 231–297 (1978).
14. Weissman, J.S. & Kim, P.S. Reexamination of the folding of BPTI: Predominance of native intermediates. *Science* **253**, 1386–1393 (1991).
15. Weissman, J.S. & Kim, P.S. Kinetic role of nonnative species in the folding of bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA* **89**, 9900–9904 (1992).
16. Brandts, J.F., Halvorson, H.R. & Brennan, M. Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* **14**, 4953–4963 (1975).
17. Hagerman, P.J. Kinetic analysis of the reversible folding reactions of small proteins: Application to the folding of lysozyme and cytochrome *c*. *Biopolymers* **16**, 731–747 (1977).
18. Schmid, F.X. & Baldwin, R.L. Acid catalysis of the formation of the slow-folding species of RNase A: Evidence that the reaction is proline isomerization. *Proc. Natl. Acad. Sci. USA* **75**, 4764–4768 (1978).
19. Kim, P.S. & Baldwin, R.L. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Ann. Rev. Biochem.* **51**, 459–489 (1982).
20. Kim, P.S. & Baldwin, R.L. Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* **59**, 631–660 (1990).
21. Mann, C.J., Shao, X. & Matthews, C.R. Characterization of the slow folding reactions of *trp* aporepressor from *Escherichia coli* by mutational analysis of prolines and catalysis by a peptidyl-prolyl isomerase. *Biochemistry* **34**, 14573–14580 (1995).
22. Englander, S.W. & Poulsen A. Hydrogen-tritium exchange of the random chain polypeptide. *Biopolymers* **7**, 329–393 (1969).
23. Woodward, C.K. & Rosenberg, A. Studies of hydrogen exchange in proteins. VI. Urea effects on RNase hydrogen exchange kinetics leading to a general model for hydrogen exchange from folded proteins. *J. Biol. Chem.* **246**, 4114–4121 (1971).
24. Englander, S.W., Downer, N.W., & Teitelbaum, H. Hydrogen exchange. *Annu. Rev. Biochem.* **41**, 903–924 (1972).
25. Woodward, C.K. & Hilton, B.D. Hydrogen exchange kinetics and internal motions in proteins and nucleic acids. *Annu. Rev. Biophys. Bioeng.* **8**, 99–127 (1979).
26. Bai, Y., Sosnick, T.R., Mayne, L. & Englander, S.W. Protein folding intermediates: Native-

state hydrogen exchange. *Science* **269**, 192–197 (1995).
27. Bai, Y. & Englander, S.W. Future directions in folding: The multi-state nature of protein structure. *Proteins: Struct. Funct. Genet.* **24**, 145–151 (1996).
28. Balbach, J. *et al*. Following protein folding in real time using NMR spectroscopy. *Nature Struct. Biol.* **2**, 865–870 (1995).
29. Udgaonkar, J.B. & Baldwin, R.L. NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature* **335**, 694–699 (1988).
30. Roder, H., Elöve, G.A. & Englander, S.W. Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature* **335**, 700–704 (1988).
31. Radford, S.E., Dobson, C.M. & Evans, P.A. The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature* **358**, 302–307 (1992).
32. Miranker, A., Robinson, C.V., Radford, S.E., Aplin, R.T. & Dobson, C.M. Detection of transient protein folding populations by mass spectrometry. *Science* **262**, 896–900 (1993).
33. Radford, S.E. & Dobson, C.M. Insights into protein folding using physical techniques: Studies of lysozyme and α-lactalbumin. *Phil. Trans. R. Soc. Lond.* B**348**, 17–25 (1995).
34. Miranker, A., Robinson, C. V., Radford, S. E. & Dobson, C. M. Investigation of protein folding by mass spectrometry. *Faseb J.* **10**, 93–101 (1996).
35. Chen, B.-L., Baase, W. A. & Schellman, J. A. Low-temperature unfolding of a mutant of phage T4 lysozyme. 2. Kinetic investigations. *Biochemistry* **28**, 691–699 (1989).
36. Matouschek, A., Kellis, J.T. Jr., Serrano, L., Bycroft, M. & Fersht, A.R. Transient folding intermediates characterized by protein engineering. *Nature* **346**, 440–445 (1990).
37. Fersht, A.R. Characterizing transition states in protein folding: An essential step in the puzzle. *Curr. Opin. Struct. Biol.* **5**, 79–84 (1995).
38. Neira, J.L. *et al*. Towards the complete structural characterization of a protein folding pathway: the structures of the denatured, transition and native states for the association/folding of two complementary fragments of cleaved chymotrypsin inhibitor 2. Direct evidence for a nucleation-condensation mechanism. *Folding & Design* **1**, 189–208 (1996).
39. Jones, C. M. *et al*. Fast events in protein folding initiated by nanosecond laser photolysis. *Proc. Natl. Acad. Sci. USA* **90**, 11860–11864 (1993).
40. Williams, S. et al. Fast events in protein folding: Helix melting and formation in a small peptide. *Biochemistry* **35**, 691–697 (1996).
41. Pascher, T., Chesick, J.P., Winkler, J.R. & Gray, H.B. Protein folding triggered by electron transfer. *Science* **271**, 1558–1560 (1996).
42. Miranker, A.D. & Dobson, C.M. Collapse and cooperativity in protein folding. *Curr. Opin. Struct. Biol.* **6**, 31–42 (1996).
43. Jackson, S.E. & Fersht, A.R. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428–10435 (1991).
44. Fersht, A.R. Optimization of rates of protein folding: The nucleation-condensation mechanism and its implications. *Proc. Natl. Acad. Sci. USA* **92**, 10869–10873 (1995).
45. Sosnick, T.R., Mayne, L., Hiller, R. & Englander, S.W. The barriers in protein folding. *Nature Struct. Biol.* **1**, 149–156 (1994).
46. Huang, G.S. & Oas, T.G. Submillisecond folding of monomeric λ repressor. *Proc. Natl. Acad. Sci. USA* **92**, 6878–6882 (1995).
47. Schindler, T., Herrler, M., Marahiel, M.A. & Schmid. F.X. Extremely rapid protein folding in the absence of intermediates. *Nature Struct. Biol.* **2**, 663–673 (1995).
48. Sosnick, T.R., Mayne, L. & Englander, S.W. Molecular collapse: The rate-limiting step in two-state cytochrome c folding. *Proteins: Struct. Funct. Genet.* **24**, 413–426 (1996).
49. Chan, H.S. & Dill, K.A. Polymer principles in protein structure and stability. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 447–490 (1991).
50. Chan, H.S. & Dill, K.A. Comparing folding codes for proteins and polymers. *Proteins: Struct. Funct. Genet.* **24**, 335–344 (1996).
51. Bryngelson, J.D., Onuchic, J.N., Socci, N.D. & Wolynes, P.G. Funnels, pathways and the energy landscape of protein folding: A synthesis. *Proteins: Struct. Funct. Genet.* **21**, 167–195 (1995).
52. Wolynes, P. G., Onuchic, J. N. & Thirumalai, D. Navigating the folding routes. *Science* **267**, 1619–1620 (1995).
53. Onuchic, J.N., Wolynes, P.G., Luthey-Schulten, Z. & Socci, N.D. Towards an outline of the topography of a realistic protein folding funnel. *Proc. Natl. Acad. Sci. USA* **92**, 3626–3630 (1995).
54. Socci, N.D., Onuchic, J.N. & Wolynes, P.G. Diffusive dynamics of the reaction coordinate for protein folding funnels. *J. Chem. Phys.* **104**, 5860–5868 (1996).
55. Dill, K.A. The stabilities of globular proteins. In Protein Engineering (eds Oxender, D. L. & Fox, C. F.) 187–192 (Alan R. Liss, Inc., New York, 1987).
56. Bryngelson, J.D. & Wolynes, P.G. Intermediates and barrier crossing in a random energy model (with applications to protein folding). *J. Phys. Chem.* **93**, 6902–6915 (1989).
57. Shakhnovich, E.I., Farztdinov, G., Gutin, A.M. & Karplus, M. Protein folding bottlenecks: A lattice Monte Carlo simulation. *Phys. Rev. Lett.* **67**, 1665–1668 (1991).
58. Camacho, C.J. & Thirumalai, D. Kinetics and thermodynamics of folding in model proteins. *Proc. Natl. Acad. Sci. USA* **90**, 6369–6372 (1993).
59. Chan, H.S. & Dill, K.A. Transition states and folding dynamics of proteins and heteropolymers. *J. Chem. Phys.* **100**, 9238–9257 (1994).
60. Socci, N.D. & Onuchic, J.N. Folding kinetics of protein-like heteropolymers. *J. Chem. Phys.* **101**, 1519–1528 (1994).
61. Abkevich, V.I., Gutin, A.M. & Shakhnovich, E.I. Free energy landscape for protein folding kinetics: Intermediates, traps, and multiple pathways in theory and lattice model simulations. *J. Chem. Phys.* **101**, 6052–6062 (1994).
62. Šali, A., Shakhnovich, E. & Karplus, M. How does a protein fold? *Nature* **369**, 248–251 (1994).
63. Dill, K.A. *et al*. Principles of protein folding—A perspective from simple exact models. *Protein Sci.* **4**, 561–602 (1995).
64. Hao, M.-H. & Scheraga, H.A. Statistical thermodynamics of protein folding: Sequence dependence. *J. Phys. Chem.* **98**, 9882–9893 (1994).
65. Hao, M.-H. & Scheraga, H.A. How optimization of potential functions affects protein folding. *Proc. Natl. Acad. Sci. USA* **93**, 4984–4989 (1996).
66. Bryngelson, J.D. & Wolynes, P.G. Spin-glass and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. USA* **84**, 7524–7528 (1987).
67. Leopold, P.E., Montal, M. & Onuchic, J.N. Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proc. Natl. Acad. Sci. USA* **89**, 8721–8725 (1992).
68. Guo, Z. & Thirumalai, D. Kinetics of protein folding: Nucleation mechanism, time scales, and pathways. *Biopolymers* **36**, 83–102 (1995).
69. Socci, N.D. & Onuchic, J.N. Kinetic and thermodynamic analysis of proteinlike heteropolymers: Monte Carlo histogram technique. *J. Chem. Phys.* **103**, 4732–4744 (1995).
70. Thirumalai, D. From minimal models to real proteins: Time scales for protein folding kinetics. *J. Phys. I* **5**, 1457–1467 (1995).
71. Mirny, L.A., Abkevich, V. & Shakhnovich, E.I. Universality and diversity of the protein folding scenarios: A comprehensive analysis with the aid of a lattice model. *Folding & Design* **1**, 103–116 (1996).
72. Thirumalai, D. & Woodson, S.A. Kinetics of folding of proteins and RNA. *Acc. Chem. Res.* **29**, 433–439 (1996).
73. Chan, H.S. & Dill, K.A. Protein folding kinetics from the perspectives of simple models. *Proteins: Struct. Funct. Genet.* **in the press**.
74. Chan, H.S. & Dill, K.A. Energy landscapes and the collapse dynamics of homopolymers. *J. Chem. Phys.* **99**, 2116–2127 (1993).
75. Kuroda, Y., Hamada, D., Tanaka, T. & Goto, Y. High helicity of peptide fragments corresponding to β-strand regions of β-lactoglobulin observed by 2D-NMR spectroscopy. *Folding & Design* **1**, 243–251 (1996).
76. Hamada, D., Segawa, S.-I. & Goto, Y. Non-native α-helical intermediate in the refolding of β-lactoglobulin, a predominantly β-sheet protein. *Nature Struct. Biol.* **3**, 868–873(1996).
77. Landry, S.J. & Gierasch, L.M. Polypeptide interactions with molecular chaperones and their relationship to *in vivo* protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **23**, 645–669 (1994).
78. Hlodan, R. & Hartl, F.U. in Mechanisms of Protein Folding (ed. R.H. Pain) 194–228 (Oxford University Press, New York, 1994).
79. Thirumalai, D. in Statistical Mechanics, Protein Structure, and Protein-Substrate Interactions (ed. S. Doniach) 115–134 (Plenum, New York, 1994).
80. Chan, H.S. & Dill, K.A. A simple model of chaperonin-mediated protein folding. *Proteins: Struct. Funct. Genet.* **24**, 345–351 (1996).
81. Todd, M.J., Lorimer, G.H. & Thirumalai, D. Chaperonin-facilitated protein folding: Optimization of rate and yield by an iterative annealing mechanism. *Proc. Natl. Acad. Sci. USA* **93**, 4030–4035 (1996).
82. Sfatos, C.D., Gutin, A.M., Abkevich, V.I. & Shakhnovich, E.I. Simulations of chaperone-assisted folding. *Biochemistry* **35**, 334–339 (1996).
83. Gutin, A.M., Abkevich, V. I. & Shakhnovich, E.I. Is burst hydrophobic collapse necessary for protein folding? *Biochemistry* **34**, 3066–3076 (1995).
84. Shrivastava, I., Vishveshwara, S., Cieplak, M., Maritan, A. & Banavar, J.R. Lattice model for rapidly folding protein-like heteropolymers. *Proc. Natl. Acad. Sci. USA* **92**, 9206–9209 (1995).
85. Klimov, D.K. & Thirumalai, D. A criterion that determines the foldability of proteins. *Phys. Rev. Lett.* **76**, 4070–4073 (1996).
86. Klimov, D.K. & Thirumalai, D. Factors governing the foldability of proteins. *Proteins: Struct. Funct. Genet.* **In the press**.
87. Hill, T.L. Effect of rotation on the diffusion-controlled rate of ligand-protein association. *Proc. Natl. Acad. Sci. USA* **72**, 4918–4922 (1975).
88. Hill, T. L. Diffusion frequency factors in some simple examples of transition-state rate theory. *Proc. Natl. Acad. Sci. USA* **73**, 679–683 (1976).
89. Steinfeld, J.I., Francisco, J.S. & Hase, W.L. Chemical Kinetics and Dynamics (Prentice-Hall, Englewood Cliffs, New Jersey, 1989).
90. Zhou, H.-X. & Zwanzig, R. A rate process with an entropy barrier. *J. Chem. Phys.* **94**, 6147–6152 (1991).
91. Tanford, C. Protein denaturation. *Adv. Protein Chem.* **23**, 121–282 (1968).
92. Matthews, C.R., Crisanti, M.M., Manz, J.T. & Gepner, G.L. Effects of a single amino acid substitution on the folding of the α-subunit of tryptophan synthase. *Biochemistry* **22**, 1445–1452 (1983).
93. Segawa, S.-I. & Sugihara, M. Characterization of the transition state of lysozyme unfolding. I. Effect of protein-solvent interactions on the transition state. *Biopolymers* **23**, 2473–2488 (1984).
94. Chen, B.-L., Baase, W.A., Nicholson, H. & Schellman, J.A. Folding kinetics of T4 lysozyme and nine mutants at 12 °C. *Biochemistry* **31**, 1464–1476 (1992).
95. Shortle, D., Chan, H.S. & Dill, K.A. Modeling the effects of mutations on the denatured states of proteins. *Protein Sci.* **1**, 201–215 (1992).
96. Chan, H.S. Kinetics of protein folding. *Nature* **373**, 664–665 (1995).
97. Unger, R. & Moult, J. Local interactions dominate folding in a simple protein model. *J. Mol. Biol.* **259**, 988–994 (1996).
98. Fukugita, M., Lancaster, D. & Mitchard, M.G. A heteropolymer model study for the mechanism of protein folding. *Biopolymers*, **in the press**.
99. Dill, K.A., Fiebig, K.M. & Chan, H.S. Cooperativity in protein folding kinetics. *Proc. Natl. Acad. Sci. USA* **90**, 1942–1946 (1993).
100. Abkevich, V.I., Gutin, A.M. & Shakhnovich, E.I. Specific nucleus as the transition state for protein folding: Evidence from the lattice model. *Biochemistry* **33**, 10026–10036 (1994).
101. Onuchic, J.N., Socci, N.D., Luthey-Schulten, Z. & Wolynes, P.G. Protein folding funnels: The nature of the transition state ensemble. *Folding & Design*, **in the press**.
102. Kiefhaber, T. & Baldwin, R.L. Kinetics of hydrogen bond breakage in the process of unfolding of ribonuclease A measured by pulsed hydrogen exchange. *Proc. Natl. Acad. Sci. USA* **92**, 2657–2661 (1995).
103. Hoeltzli, S.D. & Frieden, C. Stopped-flow NMR spectroscopy: Real-time unfolding studies of 6-^{19}F-tryptophan-labeled *Escherichia coli* dihydrofolate reductase. *Proc. Natl. Acad. Sci. USA* **92**, 9318–9322 (1995).
104. Hvidt, A. & Nielsen, S.O. Hydrogen exchange in proteins. *Adv. Protein Chem.* **21**, 287–386 (1966).
105. Miller, D.W. & Dill, K. A. A statistical mechanical model for hydrogen exchange in globular proteins. *Protein Sci.* **4**, 1860–1873 (1995).
106. Alonso, D.O.V. & Dill, K.A. Solvent denaturation and stabilization of globular proteins. *Biochemistry* **30**, 5974–5985 (1991).
107. Dill, K.A., Alonso, D.O.V. & Hutchinson, K. Thermal stabilities of globular proteins. *Biochemistry* **28**, 5439–5449 (1989).
108. Hagen, S.J., Hofrichter, J., Szabo, A. & Eaton, W.A. Diffusion-limited contact formation in unfolded cytochrom *c*: Estimating the maximum rate of protein folding. *Proc. Natl. Acad. Sci. USA* **93**, 11615–11617 (1996).
109. McCammon, J.A. A speed limit for protein folding. *Proc. Natl. Acad. Sci. USA* **93**, 11426–11427 (1996).