Protein Science (2000), *9*:10–19. Cambridge University Press. Printed in the USA.

REVIEW

Folding and binding cascades: Dynamic landscapes and population shifts

SANDEEP KUMAR,[1] BUYONG MA,[1] CHUNG-JUNG TSAI,[2] NEETI SINHA,[2] AND RUTH NUSSINOV[2,3]

[1]Laboratory of Experimental and Computational Biology, NCI-FCRDC, Bldg. 469, Rm. 151, Frederick, Maryland 21702
[2]Intramural Research Support Program—SAIC, Laboratory of Experimental and Computational Biology, NCI-FCRDC, Bldg. 469, Rm. 151, Frederick, Maryland 21702
[3]Sackler Institute of Molecular Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

(RECEIVED September 23, 1999; FINAL REVISION October 29, 1999; ACCEPTED November 5, 1999)

Abstract

Whereas previously we have successfully utilized the folding funnels concept to rationalize binding mechanisms (Ma B, Kumar S, Tsai CJ, Nussinov R, 1999, *Protein Eng 12*:713–720) and to describe binding (Tsai CJ, Kumar S, Ma B, Nussinov R, 1999, *Protein Sci 8*:1181–1190), here we further extend the concept of folding funnels, illustrating its utility in explaining enzyme pathways, multimolecular associations, and allostery. This extension is based on the recognition that funnels are not stationary; rather, they are *dynamic*, depending on the physical or binding conditions (Tsai CJ, Ma B, Nussinov R, 1999, *Proc Natl Acad Sci USA 96*:9970–9972). Different binding states change the surrounding environment of proteins. The changed environment is in turn expressed in shifted energy landscapes, with different shapes and distributions of populations of conformers. Hence, the function of a protein and its properties are not only decided by the static folded three-dimensional structure; they are determined by the distribution of its conformational substates, and in particular, by the *redistributions* of the populations under different environments. That is, protein function derives from its dynamic energy landscape, caused by changes in its surroundings.

Keywords: allostery; binding; biological pathways; conformational ensembles; dynamic landscapes; folding; funnels; induced conformational change

Biological processes are carried out through binding. Transmitting a signal, or forming an active molecular species, is the outcome of a series of binding events. Pathways, molecular communication, regulation, turning on and off genes, are all the outcome of a cascade of binding events. The essence of regulation is such that these events are often carried out in an ordered manner. Hence, from the structural standpoint, comparison of the free, unbound molecule, to the molecule when bound in a complex to a ligand, to an effector, or to another chain, frequently illustrates their variabilities. Furthermore, such conformational differences may be observed between structures as the molecules undergo through step-by-step binding events. The conformations may differ depending on whether this is the first binding event in the pathway, or in the multimolecular assembly, or is the second binding event, and so on.

Reprint requests to: Ruth Nussinov, NCI-FCRDC Bldg. 469, Rm. 151, Frederick, Maryland 21702; e-mail: ruthn@ncifcrf.gov.

Considerable research effort has gone into an investigation of the nature of these conformational differences and into their mechanisms. The general notion has been of *induced* conformational change, triggered by the binding event(s). This notion dates back to the original Koshland (1958) model. In considering biological pathways, the traditional explanation has been that a molecule undergoes a series of conformational changes, each *triggered* in turn as the molecule progresses through its sequential associations. This, however, is not the case. Rather, what we observe is a question of a shift in populations. Indeed, the one-sided view of "induced fit" is not what Koshland had proposed originally. In his seminal paper of 1958, he already depicted the enzyme as present in a range of different conformational states. Unfortunately, except for a few studies (Herschlag, 1988), what is inherently implied in the present oversimplified textbook language is that it is the binding process itself which drives the change, by optimizing the interactions between the two binding molecules.

Despite the inconsistency of this view with current beliefs regarding protein folding and binding mechanisms, this traditional concept is still widely held. Current notions of protein folding, as embodied in the funnel concept, hold that multiple conformations

glide down the slopes of the funnel through multiple routes (Bryngelson & Wolynes, 1989; Frauenfelder et al., 1991; Karplus & Shakhnovitch, 1992; Baldwin, 1994; 1995; Bozko & Brooks, 1995; Karplus et al., 1995; Onuchic et al., 1995, 1996; Wolynes et al., 1995; Dill & Chan, 1997; Karplus, 1997; Lazaridis & Karplus, 1997; Gruebele & Wolynes, 1998; Dill, 1999). Around the bottom of the funnel, there is a range of conformational isomers. The bottom of the funnel is rugged, with the conformational ensemble depending on the extent of the ruggedness, on the depths of the wells, and on the heights of the barriers separating them (Kumar et al., 1999; Ma et al., 1999; Tsai et al., 1999a). The more flexible the molecule, the larger the ensemble of conformers. Not all conformations are equally represented in solution. Some have high population times, while the probabilities for others might be very low. This ensemble represents the repertoire of molecules available for the binding event. As folding is a hierarchical process (Plaxco et al., 1998; Tsai et al., 1998, 1999a, 1999c; Baldwin & Rose, 1999a, 1999b; Sinclair & Shortle, 1999; Kragelund et al., 1999), binding and folding are similar in character (Tsai et al., 1998, 1999a, 1999c). Consequently, binding funnels have all the hallmarks of folding funnels (Tsai et al., 1999a): multiple conformations of the already folded chains race down the binding funnel slopes, toward its bottom. As in folding, the bottom of the binding funnel is similarly rugged, populated by an *ensemble of complexed* conformational isomers. This ruggedness illustrates the range of ways in which the variable conformers can bind (Kumar et al., 1999; Ma et al., 1999). Hence, folding funnels are not just an abstract concept, as has sometimes been referred to in the general biological community. Here we show the utility of the concept, through its extension to cover biological processes: when applied to hierarchical folding, multimolecular complexes, cascades of binding events, and regulation, we are able to straightforwardly rationalize these without the need to resort to the "classical" mechanisms of propagation of binding effects in the molecule.

The key to the extension resides in the recognition of the effect of the environment (Tsai et al., 1999b). Changes in the environment can be purely physical, such as pH, temperature, ionic concentration, or pressure. Alternatively, changes can also be the outcome of the binding state of the protein, that is, whether it is unbound, or bound to one, two, three, or more molecules. Hence, while in general the free energy landscape is depicted by a funnel-like shape, the details of the landscape surface of a folding funnel will differ, depending on the surroundings of the protein. A recent nice example showing the effect of a change in the physical environment is the work by Sabelko et al. (1999). These authors have proposed that the change in the free energy surface as the temperature is raised from 2 to 8 °C accounts for the observed nonexponential kinetics in the formation of the native structure. The recent elegant work by Freire and his colleagues (Freire, 1999; Todd & Freire, 1999), illustrating that residues remote from the active site are stabilized by intermolecular association, provide an example of a shift in the energy landscape of a folding funnel, caused by a binding event (Tsai et al., 1999b).

The effects of the environment on the shape of the funnels can be expressed in the funnels' walls, steepness, traps encountered by the down-racing conformations, bumps, and barrier heights. Additionally, environmental effects can also be observed in the shape of the funnels around their bottoms. Comparisons of the funnels of the chains as they undergo successive binding events, as in the case of formation of multimolecular complexes, would illustrate that as the protein rigidifies through stepwise binding events, the funnels would become progressively steeper, sharper, with narrower bottoms. Comparisons of the ensembles around their bottoms would illustrate that the distributions of the populations differ. The most prevalent conformers around the bottom of one funnel are not necessarily those that are most frequent in successive funnel(s) of a given chain. Different distributions would also be observed in allosterically regulated proteins. While the existence of different binding modes has long been recognized, their consequences, in terms of redistributions of the populations of conformers of the bound chains with respect to their unbound state, have not been considered. The redistributions explain the observed successive binding events, regulation, and hence protein function.

Here, we describe the concept of *dynamic energy landscapes*. While previously funnels have been depicted as *stationary*, considerations of the physical environment and their bound (unbound) state illustrate the shifts in the energy landscape that may take place. *Dynamic* landscapes, with altered funnel shapes and probabilities of the populations, show how the concept of the folding funnels can be utilized toward rationalization of how proteins carry out their roles in the cell. This concept of a dynamic (i.e., changing with environment and binding state) landscape has been proposed previously in the first statistical model of how chaperones can work to help the normal funneling of proteins to their native states (Gulukota & Wolynes, 1994). The distinction between static and dynamic energy landscapes is not trivial. Protein function is often inferred from single stable conformations. On the other hand, here we argue that function derives from the way that the protein is able to adjust the populations of its conformational substates as it undergoes through a series of binding events, or through different physical conditions.

It is important to distinguish between the "dynamics" term employed in molecular dynamics simulations, and the dynamic energy landscape proposed here. In molecular dynamics, we inspect snapshots of dynamically changing conformations simulated under sets of constant conditions. On the other hand, here the changes in the landscape are the outcome of changes in the environment, whether physical (e.g., temperature, pH, the presence of a denaturant) or functional (e.g., binding to other molecules).

The basic concept and its application to protein function and to biological processes

For simplicity and clarity, here we enumerate the major points embodied in the concept of dynamic landscapes and population shifts in folding and binding. The description is enhanced by two schematic diagrams (Figs. 1, 2). In the following sections of the paper, each point is explained, with the biological examples presented and classified in terms of the categories summarized in this section.

(1) In solution, there is a range of conformational isomers, in equilibrium with each other. The kinetic barriers between them are low, resulting in low energy transitions. The conformations that reside in their shallow energy wells frequently flip between the wells. Inspection of the energy landscape of Fig. 1A straightforwardly rationalizes why there is no need to resort to conformational changes *induced* by binding.

(2) There is a change in the equilibrium, the outcome of the changes in the conditions. Here, the changes in the conditions refer to binding events, such as binding to other proteins, to peptides,

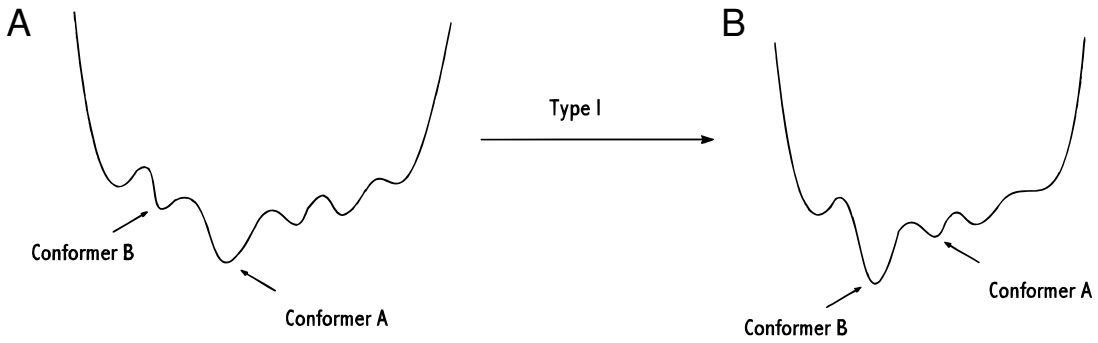

Fig. 1. A schematic drawing illustrating type I energy landscape change due to a change in the conditions of the environment. The change in the environment makes the bottom of the funnel-shape landscape shift to a different conformation. As implied by a population shift, and in contrast to the *induced-fit* mechanism, here the roughness of the landscape surface is such that no surface bump separating the conformations is substantial. This type of environmental change creates an equilibrium shift from one conformation to the other. Conformation A is the most highly populated conformer in the unbound state. Conformer B is the most highly populated conformer in the bound state. Type I landscape is observed in enzymes showing "induced-fit," in allosterically regulated enzymes, in biological pathways, and in multimolecular complexes involved in a variety of biological processes.

to DNA, RNA, and to cofactors. This situation is depicted by going from Figure 1A → Figure 1B. Biological processes falling into this category are referred to in the text below as *type I*. Enzymes showing "induced fit," allosteric regulation, enzyme pathways, biological cascades, and formation of multimolecular complexes, all fall into the type I category. Furthermore, as binding and folding are similar events, the binding of building blocks during protein folding also fall into this type I category.

(3) The kinetic barriers are high. This situation is depicted in Figure 2A. Such a situation is observed in amyloid formation, or in complex protein folding, where the misfolded molecules are stuck for a long time in their respective energy wells. The high barriers prevent the conformers from easily flipping back and forth.

(4) The barrier heights have been lowered, enabling faster conformational changes. This situation is depicted by going from Figure 2A → Figure 2B. In amyloid formation, the change in the environment involves the presence of an amyloid seed. Here conformation "A" refers to the misfolded conformation present in the amyloid. "C" refers to the conformation of the native protein. In the case of the chaperone catalyzed folding, "A" refers to the native conformation, whereas "C" is a misfolded intermediate. Here, the change in the environment or conditions refers to the binding to the chaperone. Biological processes, where initially there are high barriers and these are subsequently lowered, are classified as belonging to *type II*.

In type I, there are practically no kinetic barriers. The relative stabilities change between the conformers, the outcome of their binding conditions. Hence, the landscape changes as the protein functions. In type II, the relative stabilities are unchanged. Rather, it is the heights of the barriers that are affected. Lowering the barriers, e.g., via binding to misfolded amyloid protofibrils, results in faster population shifts, changing the equilibrium.

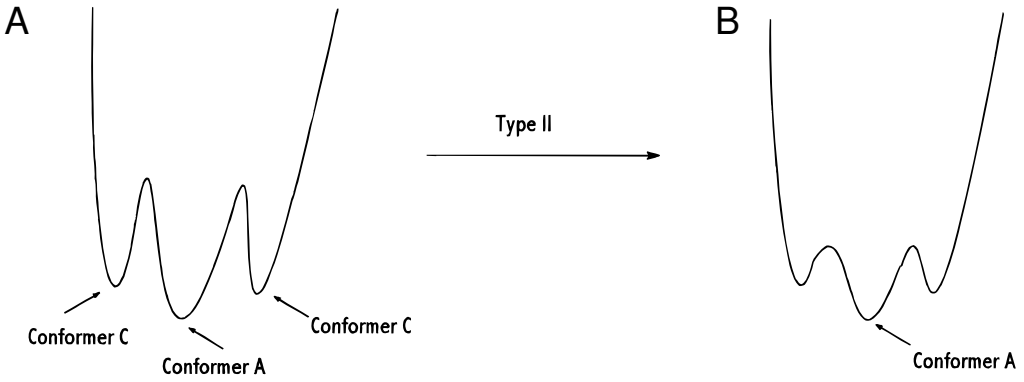

Fig. 2. A schematic drawing illustrating type II energy landscape change. The environmental change results only in an alteration of the surface roughness, but not in the relative stabilities, between a trapped conformation and the conformation at the bottom of the original landscape. The effect of the environment change is to lower the barrier bumps. Lowering the bumps enables a previously trapped conformation to climb out of its energy well, where it has resided for a long time. This type of environment change is called a kinetic shift. For the case of the amyloid formation, conformation A is the misfolded conformer. For the case of folding with the help of a chaperone, conformer A is the native conformation. Both examples are discussed in the text.

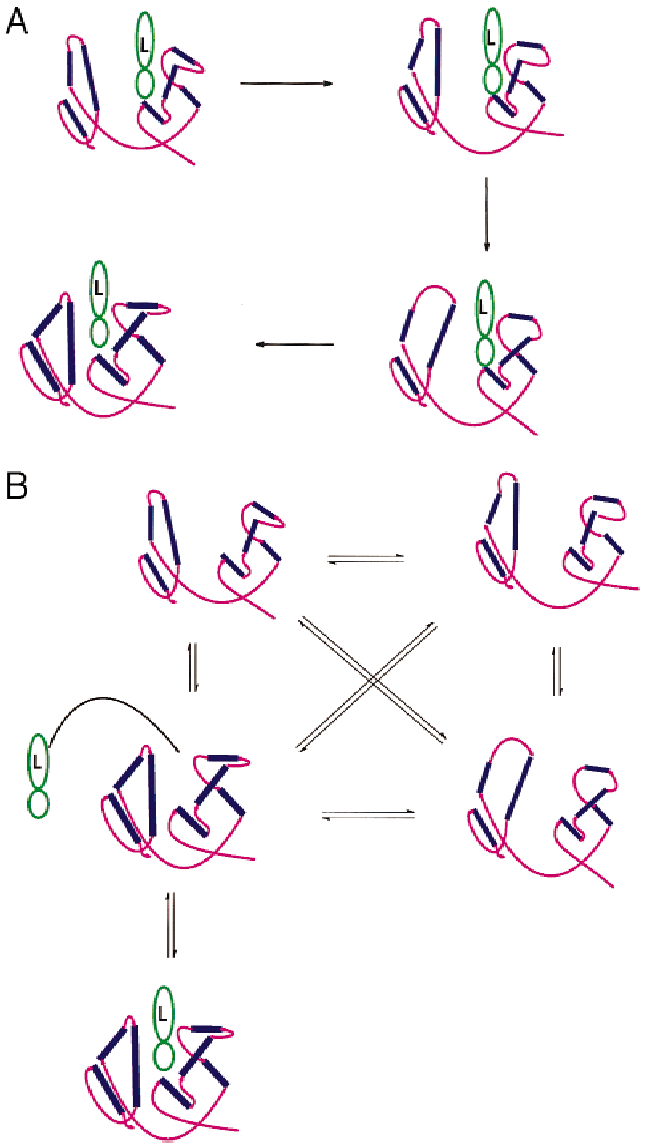

Fig. 3. An Illustration of two protein binding mechanisms. The ligand is represented by the green double circle, and the conformations of the protein chains are simplified forms of adenylate kinases from many families (Vonrhein et al., 1995). In most adenylate kinases structures, the ligands are ADP, ATP, or SO_2^{2-}. In some structures, two different ligands bind to the adenylate kinases simultaneously. In these simplified cartoons, the green double circle indicates any of the ligands or their combination. Here we propose that these conformations can coexist for a specific adenylate kinase enzyme, with different populations for different conformational substates. **A:** The traditional view of induced fit and allostery, where binding at one site causes changes at a remote site. **B:** The concept of conformational substates described here, where pre-existing conformations are in equilibrium. The equilibrium shifts to one of the conformations that fits best incoming ligand.

Below, we first describe the concept of a population in equilibrium (Fig. 3B), as an alternative for the so-called *induced fit* observation (Fig. 3A). We put this concept within the framework of a series of binding events, showing that these are all a simple, straightforward outcome of a series of sequential enrichments of certain conformations, whose binding at this time, and under these conditions, is favorable. We show how this (type I) principle holds both for the assembly of *transient building blocks* in folding and for macromolecular assemblies in binding. Both can be described under the general framework of the funnel concept: namely, *it is all a matter of dynamic shifts in the energy landscapes, during each of the binding events in hierarchical folding, in allostery, or as the molecules cascade down the biological pathway.* These are the outcome of the changing conditions or the environment (Tsai et al., 1999b).

Dynamic energy landscape and populations vs. induced conformational change

The energy landscape of proteins defines their static and dynamic properties. In Figure 1A the native conformation is around area A. At high temperatures, proteins gain enough energy and denature. For simplicity, a mountain is often used to illustrate the concept of the energy landscape and to narrate the story of folding funnels. However, in reality, the landscape is dynamic, changing with pH, ionic strength, and the presence of other molecules. The environmental change leads to a redistribution of the conformational substates. In the case of ligand binding, the most populated conformation of a protein may be different prior to, and following binding, corresponding to a change in the energy landscape.

Yet, despite the increasing understanding and realization that molecules are inherently flexible entities, existing in solution in a range of conformations (Fig. 1A), the concept of *induced fit* (illustrated in Fig. 3A) is often cited, both on a local scale, and with respect to larger, hinge-bending motions of a protein *induced* by a ligand. The argument presented here is not only that intrinsic structural flexibility facilitates the conformational transitions taking place *in response* to binding; rather, the inherent structural flexibility is responsible for the existence of a range of conformations a priori (Fig. 1A). Out of these, *pre-existing* conformations, the ones that bind are those whose structures are complementary to the molecule in question (illustrated in Fig. 3B). As the binding conformer is depleted from the solution, the equilibrium will shift in favor of this conformer, propagating the binding events (Fig. 1A $\rightarrow$ Fig. 1B). If the barriers are low, "induced fit" via population shift is directly implicated. Alternatively, it is also possible that the ligand may bind to conformers that are near the optimal one on the energy landscape. As the barriers are low, the conformational changes may take place while the ligand is already bound to the protein (Fig. 3A). Either way, the end result is the same.

This phenomenon is universal. It holds for proteins, where it is observed around the bottom of their funnels (Ma et al., 1999). It also holds for nucleic acids. DNA bending, kinking, or stretching is not *induced* by the protein. Rather, a certain fraction of DNA conformational isomers have bent, or stretched, conformations a priori. It is simply that these are the conformations whose binding to the DNA-binding proteins are more favorable than the other, unbent, conformers. Whereas their concentration when unbound is low, in the bound state their population is high. The concept also holds for the more flexible RNA conformations. Here, however, we confine ourselves to proteins.

To unambiguously demonstrate the validity of such concepts, one needs to show that the conformation, which is observed to be significantly different in the "bound" state as compared to that in the "free," unbound form, and hence, which has been considered to be the outcome of *induced fit*, also exists in the unbound form. If that conformation has a low population time, it is likely to be difficult to detect experimentally. On the other hand, it might be

feasible to capture via long-enough molecular simulations. Additionally, it will also be illuminating to examine molecules showing no apparent conformational changes between two crystals of the bound and unbound molecular species. These molecules are not necessarily rigid either. While they might appear rigid under one set of (binding) conditions, they may illustrate their conformational variability under different conditions, for example, in binding to different ligands or in different crystal forms (Muller et al., 1998).

This rationale is in agreement with a number of observations. The view is consistent with the broad range vs. specific binding. The more flexible the molecule, the larger the ensemble of conformations that it possesses. These conformations may bind a range of ligands whose structures are complementary to the ensemble. On the other hand, the more specific, selective binding is frequently attained by the more rigid molecules, which exist in fewer potential conformations. A particularly fitting example here is that of the germline, polyreactive vs. the matured, specific antibodies (Foote & Milstein, 1994).

If the binding and the subsequent reaction could be monitored, the two mechanisms (Fig. 3A vs. Fig. 3B) might be differentiated by their kinetics. The induced fit mechanism (Fig. 3A) may be expressed as

$$R + L \rightleftharpoons RL \rightleftharpoons (RL)^*. \quad (1)$$

According to this mechanism, the ligand (L) and the receptor (R) bind first *followed* by an isomerization of the complex RL to yield the activated form $(RL)^*$.

In the second mechanism (Fig. 3B), conformational shifts *predate* binding. However, the substrate can only bind to one of the enzyme's conformational states:

$$\langle R', R'', R''', \ldots \rangle + L \rightleftharpoons (RL)^* \quad (2)$$

where R', R'', R''' … are different conformers of the receptor R.

Results of kinetic experiments recorded for the glutathione transferases (Stella et al., 1998; Nieslanik et al., 1999) and for ester-hydrolyzing antibodies raised against a phosphonate transition state analogue (Geyer et al., 1996; Lindner et al., 1999), which seem to indicate an induced fit type mechanism, are equally consistent with a pre-existing equilibrium between high and low affinity conformers. Crystal structures of these antibodies, free and complexed, have indicated that their conformations do not change upon binding.

For the human glutathione transferase P1-1, Stella et al. (1998) have clearly demonstrated that in the absence of glutathione, the apoenzyme exists in at least two different families of conformational states (polar, 38%; apolar, 62%). In the presence of saturating glutathione concentration, the equilibrium is shifted toward the apolar conformers (83% of the total population). The existence of the bound conformational isomer of glutathione transferase P1-1 at this (38%) concentration while still in the unbound state is fortunate, as it is high enough to be easily detected. Unfortunately, in general this is not the case.

These "induced fit" examples belong to type I category (Fig. 1). The changing conditions that cause the change in the energy landscape (Fig. 1A → Fig. 1B) are the events of binding to the ligands, i.e., the phosphonate transition state analogue or the glutathione in the examples here. The binding event creates different surrounding environment for the receptor, making the less stable conformation "B" in the unbound state (Fig. 1A) more stable in the bound state (Fig. 1B). On the other hand, the more stable conformation "A" in the unbound state is less favorable for binding the ligand and is not selected.

Hence, there are two mechanisms, both leading to the same end result. In the conformational selection (Fig. 3B), it is not the binding itself that elicits the conformational change. Instead, the alternate conformation exists a priori. However, its population time may be low. It is also consistent with the slow process of domain swapping (Bennett et al., 1994, 1995). In this intriguing swapping phenomenon, proteins have been crystallized in two forms: in the first, two (sub)domains of the protein chain interact with each other. In the second, alternate, form, the two domains interact with sister-domains from a second monomer. The process of the swapping may be slow, on the order of hours to days (D'Alessio, 1995). Within the conformational ensemble around the bottom of the funnel, there exists the conformer having its domain swung out. Its population may be low, since in the unbound state it is likely to be considerably less stable. However, once bound, a stable complex is achieved, with the equilibrium in the ensemble shifting itself in favor of this flipped conformer. Alternatively, it is also possible that the ligand binds to an alternate conformation, which is near it at the bottom of the funnel. As the barriers are low, the conformational transition may take place in the bound state, again resulting in shifting the equilibrium in its direction (Fig. 3A).

The principles of binding are general: molecules bind to each other if their conformations are complementary, in geometry and chemistry, and their binding produces stable associations (Cherfils & Janin, 1993; Norel et al., 1995, 1999). The conformers that bind are not necessarily those whose populations are the highest in solution. Rather, those that associate are the ones whose conformations are most favorable and produce more stable complexed bound structures. These principles hold whether in single molecule crystals, or in crystals of complexed structures. If the conformations that associate have low population times in solution, owing to the equilibrium in the ensemble, their binding will result in flipping of other conformations in favor of the conformation of the bound conformer.

Binding: Populations in allostery and in biological pathways

Classical biochemistry has long taught us that proteins are the major components in biological processes. Furthermore, one of the critical ways through which proteins exert their regulatory control is *allostery*. Allosteric proteins typically have at least two separate binding sites, one for the substrate and one for the regulatory ligand. These sites are at different locations on the protein surface. Since the binding of the ligand at one site affects the conformation of the second, these proteins play a key role in a broad range of regulatory, metabolic, or other processes, regardless of their chemical nature (Alberts et al., 1989). Allosteric proteins mediate sensitive responses to signals. This has been traditionally explained through a cooperative change in the conformations of their subunits. It has been proposed that the binding of a ligand induces a conformational change, which in turn elicits a change in identical subunits in the symmetrical protein assembly (Alberts et al., 1989). If the concentration of the ligand is low, such a cooperative effect can result in a substantial speeding of the rate of the reaction. The widely popular notion is that allosteric proteins have at least two distinct conformations, separated from each other by unstable intermediate states (Alberts et al., 1989, and references therein). The molecule will "flip-flop" between the stabilized conformations, depending on the presence, or absence, of the ligands. Enzymes

early in a pathway are often allosterically regulated and can exist in active and inactive conformations. Depending on the role of the enzyme, it may illustrate a positive feedback or a negative inhibition, where it is either the presence of the substrate or of the end product that causes the enzyme to flip its conformation. Using the language of the textbook, this is carried out through "pushing" or "pulling" of the proteins into different shapes, with the energy necessary for this "pulling," and fitting of the shape of the protein supplied by the weak binding of the ligand. Not much energy is needed for this conversion, explaining the reversibility of the conformational changes between the "active" and "inactive" states upon the binding of such ligand (Alberts et al., 1989). On the other hand, many allosteric interchanges such as the GroEL/GroES molecular chaperone discussed below require the hydrolysis of ATP.

Precisely how do the changes in conformation, which occur at the binding site owing to the binding of the ligand, propagate to the location to the second binding site to cause this allosteric conformational change has not been understood. It has been proposed that this propagation can be either through the movement of the backbone or through interactions between side chains, where these travel via the compact protein interior or through some cavities or channels.

Recently, Freire (1999) has provided an insight into this problem. Freire has shown that binding of a ligand can elicit changes in the stability of residues that are remote from the binding site. Freire has further shown that binding serves to *stabilize* these regions. This result is consistent with hydrogen exchange studies, showing that three different antibodies (D44.1, D1.3 and HyHEL-5) bound to the lysozyme at different sites, changed the protection of remote residues (Williams et al., 1996). Moreover, Freire has provided an attractive explanation: he proposed that the stabilization that is observed is through a redistribution of the populations of the substates. Conformers in which the remote residues are more stable have a higher population time than those in which the corresponding residues are more flexible; that is, there is a change in the populations between the bound and the unbound states. Freire has shown it to hold for a limited local conformational change rather than for the larger global, slower mode hinge-bending motions. Nevertheless, this explanation is in agreement with our proposition (Kumar et al., 1999; Ma et al., 1999; Tsai et al., 1999a), namely, that for the most part, the populations of the conformers are there a priori. It is mostly a question of *selection* (Fig. 3B), not of induced (Fig. 3A) conformational change, as has been traditionally assumed, although the latter can also take place, if the conformers are nearby on the energy landscape. Allosterically regulated proteins illustrate a type I energy landscape (Fig. 1A,B). A conformer with a low population time around the bottom of one funnel is enriched, with a consequently higher population time observed around the bottom of the next funnel in the binding cascade.

Folding and binding are similar processes (Tsai et al., 1998). Hence, we need to inspect the populations around the bottoms of the folding and of the binding funnels. In either case, the key is in *dynamic shifts of energy landscapes* caused by the changing (binding) conditions.

Protein folding: Populations of building blocks hierarchically assemble to yield the native fold

Protein folding can be described as a combinatorial assembly of such a set of *transient, highly populated*, contiguous, *building block* fragments (Tsai et al., 1998, 1999b). A building block may be of a variable size. It is determined by the local coding of the sequence, that is, by local interactions. If the building block is isolated, its conformation is unstable. It is the mutual interactions between building blocks that stabilize its conformation. Via combinatorial assembly building blocks bind, to form a stable, higher-population time conformation, hydrophobic folding unit. The hydrophobic folding units assemble to form domains, which in turn further assemble to form subunits and oligomers. In terms of the folding funnel landscape, the entire folding/binding process may be viewed as sequentially fusing and modifying individual funnels.

Building blocks illustrate a type I energy landscape (Fig. 1). The continuous switch from one conformation to the other is enabled by the low barriers between them (Fig. 1A). Figure 4 illustrates one such example of the β-2-microglobulin (PDB code 1bmg, Bernstein et al., 1977). The binding of the building blocks to create the higher order structure (the hydrophobic folding units) changes the conditions, resulting in a shift of the energy landscape (Fig. 1A $\rightarrow$ Fig. 1B), and hence in the equilibrium. Now the most populated conformer may either be the same as in the unbound, stand-alone building block conformer, or different.

Conformational ensembles and flexibility: Shifting the energy landscape

Conformational differences have been observed for many proteins, e.g., the germline 48G7 antibody (Wedemayer et al., 1997), triglyceride lipase, triose phosphate isomerase, phosphoglycerate kinase, adenylate kinase (Fig. 3), calmodulin, glutamate dehydrogenase, and GroEL (Gerstein & Krebs, 1998). All illustrate a type I energy landscape (Fig. 1).

Hemoglobin has been among the best studied allosteric proteins (Frauenfelder et al., 1991). The binding of oxygen to hemoglobin is mediated by proton, carbon dioxide, and organic phosphates. Biologically active adult hemoglobin is a tetramer, consisting of two α- and β-chains. Hemoglobin exists in equilibrium between two alternate states. The first has low affinity for oxygen. It is constrained by salt bridges between the C-termini of the four subunits. This state is the Tense or Taut (T). The second is the relaxed (R) state. The equilibrium between the relaxed and taut states is governed primarily by the positions of the iron atoms relative to the heme porphyrin rings, attached to each subunit. The difference between the two states consists of a rotation and translation of one $\alpha\beta$ dimer relative to the other. The change in the energy landscape (Fig. 1A $\rightarrow$ Fig. 1B) between the two forms is the result of the availability of oxygen.

Energy landscape dynamics is also useful in understanding signal transduction via G proteins. Signaling molecules, such as hormones, growth factors, and neurotransmitters, bind to their cognate cell surface receptor proteins with high specificity. These surface receptors relay the signals across the plasma-membrane and produce new intracellular signals. From the structural standpoint, G-protein linked signal transduction is one of the most thoroughly studied signaling cascade. A key step in the cascades is the binding of GTP or GDP to the G protein. G proteins are a superfamily of GTP hydrolases. Available sequence and structural data suggest that all members of this group share a common ancestor and a common structural core, exemplified by that of $p21^{ras}$ (Sprang, 1997). G proteins form stable complexes with their substrate (GTP) and product (GDP). In all G proteins, binding and hydrolysis of GTP trigger reciprocal conformational changes within the catalytic

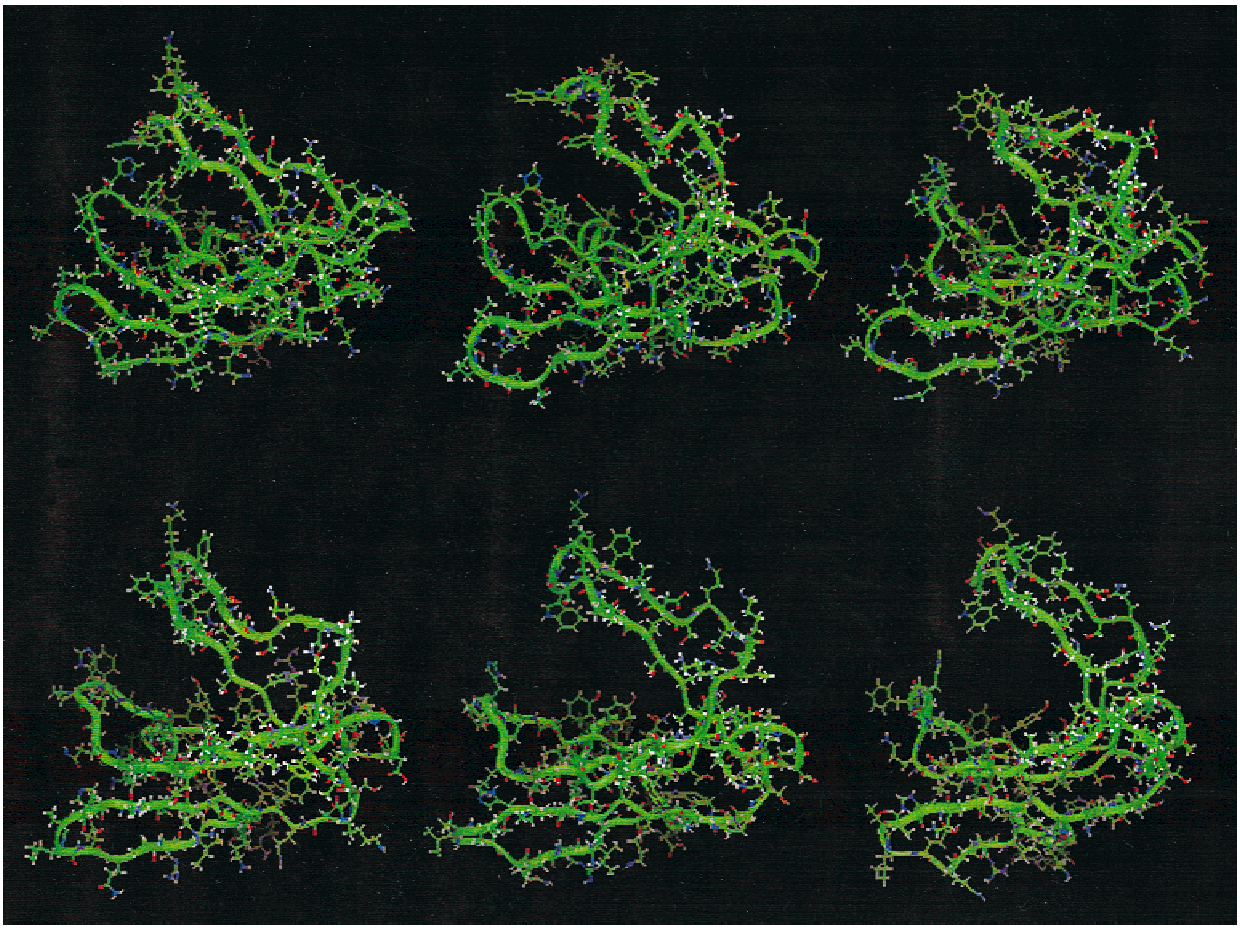

Fig. 4. An illustration of the populations of different conformations of a building block in the process of protein folding. The protein is histocompatibility antigen of bovine β-2-microglobulin (PDB code 1bmg, Bernstein et al., 1977). The conformations have been obtained from molecular dynamic simulations of high temperature protein unfolding trajectories. One building block (at the top of each conformer in the figure) is observed to open up in the unfolding process. These conformations may coexist in the folding pathway.

domain. The GTP- and GDP-bound complexes define, respectively, the active and inactive states of a G-protein as a regulator. The energy landscape of G-proteins can be considered to flip-flop between the active and inactive minima. The populations are controlled by the relative concentrations of the substrate and the product. The G proteins dynamic energy landscapes are affected by the rate of GTP hydrolysis, by binding of a GTPase activating protein, and by association with their cognate target effectors. Current data indicate that the G protein regulatory apparatus depends on a single structural element common to all G proteins: the switch II helix (Sprang, 1997), tensioned into a high energy state by GTP binding, which facilitates the binding to the target effector. Upon GTP hydrolysis, the energy is dissipated with concomitant collapse of switch II and effector release. The change in the environment leading to the Figure 1A $\rightarrow$ Figure 1B landscape shift is caused by the binding of the signaling molecule to the receptor, which in turn causes GTP to displace GDP from the G-protein, followed by the dissociation of the G-α subunit. The G-α subunit proceeds to bind adenylate kinase.

The folding energy landscape of a protein may be replete with minima deep enough to trap misfolded conformations. In a living cell, it is crucial for proteins to fold into their native conformations. The allosteric Hsp60/GroEL and TF55/CCT chaperone families are double-ring shaped complexes. The *Escherichia coli* cochaperone, GroES, caps the GroEL at one end to form a bullet-like structure, or at both ends to form a football-like shape (Sigler et al., 1998). Apparently the chaperones recognize exposed hydrophobic surfaces of a wide range of nonnative conformations and bind them in their central cavities. The allosterically-regulated, ATP to ADP mediated, cooperative motion of the seven subunits in each of the two rings, and of the two rings with respect to each other in GroEL, is the outcome of redistributions of bound and unbound conformations. Thus, this change in the populations will be observed only if the concentration of the substrate, i.e., of (misfolded) chains, is high enough. With respect to the energy landscape of the protein substrate, since these molecular chaperones catalyze protein folding by facilitating the trapped, misfolded conformations to get out of their energy minima wells; in essence, the funnel of the protein substrate is basically unchanged. On the other hand, by aiding trapped molecules, GroEL reduces the probability of misfolded conformers to reside in their wells. The allosteric switch of the chaperones is caused by the binding of the misfolded conformation, which causes the shift in the landscape of this type I protein.

Changes in the environment and dynamic energy landscapes

Changes in the environment may involve different temperature (Sabelko et al., 1999), pH, ionic strength, and pressure. Alternatively, they may include the presence in the solution of other molecules. Here we provide two examples of the effects of changes of the environment: amyloid formation and folding with the help with molecular chaperones (Gulukota & Wolynes, 1994). Both cases illustrate a type II landscape shift, as depicted in Figure 2.

The presence or absence of the amyloid fiber results in a change of the landscape of the monomeric chain. The observed shift is the effect on the barrier heights. In the absence of a preformed amyloid, the ensemble of conformations will also contain a certain concentration of misfolded monomers. However, the population of these conformers will be very low, with the corresponding funnel illustrating high barriers (Fig. 2A). However, if the solution is seeded with a pre-existing fiber, the energy landscape will change: the barriers will be lowered, and the probability of occurrence of the misfolded conformer will increase (Fig. 2B). A nice example is the polyglutamine containing huntingtin fragments that self-assemble into amyloid-like fibrils, the causative agents of Huntington disease. It has been shown that the N-terminal huntingtin fragments with polyglutamine (polyQ) tracts within the range of 51–122 residues aggregate; whereas if they are within the normal 20–30 glutamine repeat range, no aggregation is observed. Furthermore, the formation of amyloid-like aggregates is polyQ repeat length dependent (Scherzinger et al., 1999). No aggregation has been observed with polyQ length of seven residues. This suggests that there are two important factors: the nature of the proteins within the same environment—that is, the behavior of polyQ with a length of seven residues as compared to, say, 31—and the change of the environment through seeding. It is quite likely that an amyloid form of the huntingtin fragments is also present for a length of seven glutamines. However, the population of this conformer is so low that in the absence of a seed the rate of aggregation is far too slow to detect.

A second, particularly exciting example of the effect of the change of the surrounding environment on the energy landscape is that of the *proregion* molecular chaperone. The proregion is a sequence that is covalently linked to the N-termini of extracellular serine proteases, such as subtilisin, α-lytic protease and aqualysin from bacteria, and carboxypeptidase Y from yeast (Ellis, 1998, and references therein). The proregion has been termed intramolecular chaperone, as it is essential for the correct folding of the enzyme (Baker & Agard, 1994; Shinde et al., 1997; Ellis, 1998). After the chain folds to its native state, it is cleaved from the mature protein. If the proregion is removed and the enzyme is denatured, upon removal of the denaturant the chain fails to refold correctly, owing to the high barriers trapping misfolded conformers (Fig. 2A). However, if the proregion peptide is added to the solution, the barriers are lowered (Fig. 2B) and the protein folds into its native state (Ellis, 1998). Interestingly, when the proregion is mutated at a single site (isoleucine to valine, at position 48), the enzyme folds into an altered state (Shinde et al., 1997). If the proregion peptide is subsequently added to the solution, this altered conformation does not flip into the "correct," native conformation. The altered conformer has also been incubated at 4 °C for two weeks, without a change in its conformation, demonstrating its stability at this temperature range. The two conformations appear different by several measurements: CD spectra, temperature stability, threefold differences in the Michaelis constants for a synthetic substrate, and larger inhibition constants upon binding of the prosequence. Furthermore, it has been observed that the native proregion binds more tightly to the native protein fold that it helped produce, as compared to the altered conformation. Hence, here we have the same sequence that has been observed to fold into slightly different conformations, depending on its proregion environment. In the language of energy landscapes and folding funnels, the most populated enzyme conformations differ between the two cases, depending on the sequence (conformation) of the proregion peptide. Both types of conformations are present a priori. However, their distributions differ as a function of their environment. As in the case of the amyloid and of domain swapping (Bennett et al., 1994, 1995), the presence of the proregion peptide lowers the kinetic barriers, shifting the energy landscapes.

Single molecule- or molecular ensemble-chemistry: The evolutionary advantage of populational shifts over induced fit

During evolution, nature has been confronted with the necessity of choosing between two options: to optimize an induced fit mechanism or, alternatively, optimize populational shifts. On the face of it, it appears that optimizing an induced fit mechanism is more advantageous, as it would lead to an increase in binding affinity and selectivity. Nevertheless, in terms of evolution, multiple conformations is a better choice than focusing on a single conformation (Joyce, 1997). Consistent with this notion, recent evidence suggests that antigens are recognized by conformational selection (Berger et al., 1999). The evolutionary advantage of populational shifts (Joyce, 1997) has been shown by the beautiful work of crystallizing and comparing the structures of a germline antibody complex and the corresponding affinity-matured antibody (Wedemayer et al., 1997). As illustrated in Figure 3, populational shifts are far more tolerant to changes in the environment. An accidental debilitating mutational event may block a given conformational change (Fig. 3A). However, through populational shifts, there are alternate pathways to achieve a binding conformer (Fig. 3B). On the other hand, should a critical mutation take place in an induced fit pathway, it would seriously jeopardize the conformational switch channel. A parallel process, in which in principle every individual in the population has the opportunity to give rise to novel variants with increased fitness, is clearly advantageous to the organism. This parallelism makes it less likely that the population as a whole will get trapped in an evolutionary blind alley, from which further improvements in fitness are precluded (Joyce, 1997). This difference might be crucial for signal transduction. Populational shifts stand a much better chance than induced fit in ensuring signal transmission and protein function.

The structure of a protein and its properties result from natural selection. During evolution, proteins with nonrandom sequences, which are able to fold rapidly within biologically relevant time scales, have been selected. Rather than proceeding via a single, specific Levinthal (1969) folding pathway, it is far more advantageous to adopt a multiple-route folding funnel. Consistently, as folding and binding are sequential, integrate steps to achieve an optimal, robust functioning protein, we observe a multiple-way populational shift. Thus, as the general opinion holds, the properties of proteins are the outcome of the static folded three-dimensional structures and of the distribution of their conformational substates. However, we propose that, in particular, the properties derive from

the *redistributions* of the substate populations under different environments. And in biological processes and pathways, the environment undergoes constant changes.

Conclusions

The function of a molecule is mediated through its binding. Apart from the binding of inhibitors, proteins exert their effect by being part of biological processes. A protein may have built-in regulatory capabilities, such as in the case of allostery, or catalyze reactions at some advanced steps in cascading pathways. Frequently, protein molecules function by being components in a multimolecular assembly.

Examination of the conformational states of the proteins when in the bound vs. the unbound state, or when bound to different ligands/effectors, as in the case of allostery, has nurtured the view that the conformational change is elicited by the binding, with subsequent propagation through the structure (Alberts et al., 1989). It has been proposed that even a weak binding ligand produces an energy large enough to enable this conformational change. The widely accepted mechanism for a cooperative allosteric conversion in symmetric multisubunit proteins is that upon binding of the ligand to one subunit, the induced conformational change propagates to the other subunits. While the binding of the first regulatory ligand to the enzyme is a slow step, there is a rapid binding of additional regulatory ligands to the remaining subunits. Thus, the relative enzyme activity for multisubunit allosteric enzymes increases much more steeply than observed for a single subunit enzyme. Similarly, the orderly formation of multimolecular complexes, essential for the living cell as they increase the level of the local concentration of the ligands, has been attributed to induced and propagated conformational changes.

On the other hand, here we propose that these can be straightforwardly rationalized by the consequences of dynamic changes in the energy landscapes, and hence in populations. The first step of the regulatory ligand binding is slow since the population of the conformer whose binding to the ligand is favorable is low. In the bound state the population of this conformer is substantially higher. This shift in the population will be reflected in the more rapid binding observed in subsequent steps. Hence, rates are a function of populations.

The same principles apply in folding: in folding, the transient building blocks associate, forming the independently folding, hydrophobic unit. If the conformation of the building block when in the bound state is similar to the one that the corresponding peptide would have in solution, hydrophobic collapse would be fast. If, on the other hand, the conformation in the native state differs from that of the peptide in solution, binding would be considerably slower, as the population of the corresponding, binding conformer, is low. Hence, the critical issue is the barrier heights: if the conformations of the native and the peptide-building block are similar, the barriers are low, and a fast two state-like folding is observed.

The key is the dynamic changes in the energy landscapes, manifested in the dynamic shifts in the populations in subsequent steps down the cascade. This applies to the populations around the bottoms of the building blocks microfunnels in folding and to the funnels of the hydrophobic folding units, domains, and subunits. It applies equally well to biological regulation, allostery, and pathways. It is the barrier heights and the populations that change with the binding cascade; it is not the conformational changes induced and propagated by "pulling" or "pushing." Additionally, the changes of the populations, and of the funnel shapes, are the outcome of changes in the environment, whether physical or binding-state. Funnels' walls and bottoms are not stationary; energy landscapes of a given chain are dynamic, depending on the conditions.

Conformational selection is not a newly dressed-up "lock and key" mechanism. Conformational selection does not strictly imply that a specific protein conformation matches precisely a specific substrate conformation. Such a "lock and key"-like mechanism will put us at the other extreme of a narrow interpretation. Here we propose redistributions of protein conformations under different environments. In the process of such redistributions, a certain small extent of "induced fit" may still be operational locally.

Acknowledgments

We thank Dr. Jacob V. Maizel for many helpful and insightful discussions. The research of R. Nussinov in Israel has been supported in part by grant number 95-00208 from BSF, Israel, by a grant from the Ministry of Science, by the Center of Excellence, administered by the Israel Academy of Sciences, by the Magnet grant, and by the Tel Aviv University Basic Research and Adams Brain Center grants. This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under contract number NO1-CO-56000. The content of this publication does not necessarily reflect the view or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organization imply endorsement by the U.S. Government.

References

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD. 1989. *Molecular biology of the cell*, 2nd ed. New York, London: Gatland Publishing Inc.

Baker D, Agard DA. 1994. Kinetics versus thermodynamics in protein folding. *Biochemistry 33*:7505–7509.

Baldwin RL. 1994. Matching speed and stability. *Nature 369*:183–184.

Baldwin RL. 1995. The nature of protein folding pathways: The classical versus the new view. *J Biomol NMR 5*:103–109.

Baldwin RL, Rose GD. 1999a. Is protein folding hierarchic? I. Local structure and peptide folding. *Trends Biochem Sci 24*:26–33.

Baldwin RL, Rose GD. 1999b. Is protein folding hierarchic? II. Folding intermediates and transition states. *Trends Biochem Sci 24*:77–84.

Bennett MJ, Choe S, Eisenberg D. 1994. Domain swapping: Entangling alliances between proteins. *Proc Natl Acad Sci USA 91*:3127–3131.

Bennett MJ, Schlunegger MP, Eisenberg D. 1995. 3D domain swapping: A mechanism for oligomer assembly. *Protein Sci 4*:2455–2468.

Berger C, Weber-Bornhauser S, Eggenberger J, Hanes J, Pluckthun A, Bosshard HR. 1999. Antigen recognition by conformational selection. *FEBS Lett 450*:149–153.

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Bozko EM, Brooks CL III. 1995. First principles calculation of the folding free energy for a three helix bundle protein. *Science 269*:393–396.

Bryngelson JD, Wolynes PG. 1989. Intermediates and barrier crossing in a random energy model (with applications to protein folding). *J Phys Chem 93*:6902–6915.

Cherfils J, Janin J. 1993. Protein docking algorithms: Simulating molecular recognition. *Curr Opin Struct Biol 3*:265–269.

D'Alessio G. 1995. Oligomer evolution in action. *Nat Struct Biol 2*:11–13.

Dill KA. 1999. Polymer principles and protein folding. *Protein Sci 8*:1166–1180.

Dill KA, Chan HS. 1997. From Levinthal to pathways to funnels. *Nat Struct Biol 4*:10–19.

Ellis RJ. 1998. Steric chaperones. *Trends Biochem Sci 23*:43–45.

Foote J, Milstein C. 1994. Conformational isomerism and the diversity of antibodies. *Proc Natl Acad Sci USA 91*:10370–10374.

Frauenfelder H, Sligar SG, Wolynes PG. 1991. The energy landscapes and motions of proteins. *Science 254*:1598–1603.

Freire E. 1999. The propagation of binding interactions to remote sites in proteins: Analysis of the binding of the monoclonal antibody D1.3 to lysozyme. *Proc Natl Acad Sci USA 96*:10118–10122.

Gerstein M, Krebs W. 1998. A database of macromolecular motions. *Nucl Acids Res 26*:4280–4290.

Geyer M, Schweins T, Herrmann C, Prisner T, Wittinghofer A, Kalbitzer HS. 1996. Conformational transitions in p21^{ras} and in its complexes with the effector protein raf-RBD and the GTPase activating protein GAP. *Biochemistry 35*:10308–10320.

Gruebele M, Wolynes P. 1998. Satisfying turns in folding transitions. *Nat Struct Biol 5*:662–665.

Gulukota K, Wolynes P. 1994. Statistical mechanics of kinetic proofreading in protein folding *in vivo*. *Proc Natl Acad Sci USA 91*:9292–9296.

Herschlag D. 1988. The role of induced fit and conformational changes of enzymes in specificity and catalysis. *Bioorg Chem 16*:62–96.

Joyce GF. 1997. Evolutionary chemistry: Getting there from here. *Science 276*:1658–1659.

Karplus M. 1997. The Levinthal paradox: Yesterday and today. *Fold Des 2*:S69–S75.

Karplus M, Sali A, Shakhnovitch E. 1995. Comment: Kinetics of protein folding. *Nature 373*:664–665.

Karplus M, Shakhnovitch EI. 1992. Protein folding: Theoretical studies of thermodynamics and dynamics. In: Creighton T, ed. *Protein folding*. New York: Freeman & Sons. pp 127–195.

Koshland DE Jr. 1958. Application of a theory of enzyme specificity to protein synthesis. *Proc Natl Acad Sci USA 44*:98–123.

Kragelund BB, Osmark P, Neergaard TB, Schiodt J, Kristiansen K, Knudsen J, Poulsen FM. 1999. The formation of a native-like structure containing eight conserved hydrophobic residues is rate limiting in two-state protein folding of ACBP. *Nat Struct Biol 6*:594–601.

Kumar S, Ma B, Tsai CJ, Wolfson H, Nussinov R. 1999. Folding funnels and conformational transitions *via* hinge-bending motions. *Cell Biochem Biophys 31*:141–164.

Lazaridis T, Karplus M. 1997. "New view" of protein folding reconciled with the old through multiple unfolding simulations. *Science 278*:1928–1931.

Levinthal C. 1969. How to fold graciously. In: Debrunner P, Tsibris JCM, Munck E, eds. *Mossbauer spectroscopy in biological systems, proceedings of a meeting held at Allerton House, Monticello, Illinois*. Urbana, IL: University of Illinois Press. p 22.

Lindner AB, Eshhar Z, Tawfik DS. 1999. Conformational changes affect binding and catalysis by ester-hydrolysing antibodies. *J Mol Biol 285*:421–430.

Ma B, Kumar S, Tsai CJ, Nussinov R. 1999. Folding funnels and binding mechanisms. *Protein Eng 12*:713–720.

Muller Y, Kelley RF, De Vos AM. 1998. Hinge bending within the cytokine receptor superfamily revealed by the 2.4 Å crystal structure of the extracellular domain of the rabbit tissue factor. *Protein Sci 7*:1106–1115.

Nieslanik BS, Dabrowski MJ, Lyon RP, Atkins WM. 1999. Stopped-flow kinetic analysis of the ligand-induced coil-helix transition in glutathione S-transferase A1-1: Evidence for a persistent denatured state. *Biochemistry 38*:6971–6980.

Norel R, Lin SL, Wolfson H, Nussinov R. 1995. Molecular surface complementarity at protein-protein interfaces: The critical role played by surface normals at well placed, sparse, points in docking. *J Mol Biol 252*:263–273.

Norel R, Petrey D, Wolfson H, Nussinov R. 1999. Examination of shape complementarity in docking of *unbound* proteins. *Proteins 36*:307–317.

Onuchic JN, Socci ND, Luthey-Schulten Z, Wolynes PG. 1996. Protein folding funnels: The nature of the transition state ensemble. *Fold Des 1*:441–450.

Onuchic JN, Wolynes PG, Luthey-Schulten Z, Socci ND. 1995. Towards an outline of the topography of a realistic protein folding funnel. *Proc Natl Acad Sci USA 92*:3626–3630.

Plaxco KW, Simons KT, Baker D. 1998. Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol 277*:985–994.

Sabelko J, Ervin J, Gruebele M. 1999. Observations of strange kinetics in protein folding. *Proc Natl Acad Sci USA 96*:6031–6036.

Scherzinger E, Sittler A, Schweiger K, Heiser V, Lurz R, Hasenbank R, Bates GP, Lehrach H, Wanker EE. 1999. Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: Implications for Huntington's disease pathology. *Proc Natl Acad Sci USA 96*:4604–4609.

Shinde UP, Liu JJ, Inoye M. 1997. Protein memory through altered folding mediated by intramolecular chaperones. *Nature 389*:520–522.

Sigler PB, Xu Z, Rye HS, Burston SG, Fenton WA, Horwich AL. 1998. Structure and function in GroEL-mediated protein folding. *Ann Rev Biochem 67*:581–608.

Sinclair JF, Shortle D. 1999. Analysis of long-range interactions in a model denatured state of staphylococcal nuclease based on correlated changes in backbone dynamics. *Protein Sci 8*:991–1000.

Sprang SR. 1997. G protein mechanisms: Insights from structural analysis. *Ann Rev Biochem 66*:639–678.

Stella L, Caccuri AM, Rosato N, Nicotra M, Bello ML, Matteis FD, Mazzetti AP, Federic G, Ricci G. 1998. Flexibility of helix 2 in the human glutathione transferase P1-1. *J Biol Chem 273*:23267–23273.

Todd MJ, Freire E. 1999. The effects of inhibitor binding on the structural stability and cooperativity of the HIV-1 protease. *Proteins 36*:147–156.

Tsai CJ, Kumar S, Ma B, Nussinov R. 1999a. Folding funnels, binding funnels and protein function. *Protein Sci 8*:1181–1190.

Tsai CJ, Ma B, Nussinov R. 1999b. Folding and binding cascades: Shifts in energy landscapes. *Proc Natl Acad Sci USA 96*:9970–9972.

Tsai CJ, Maizel JV, Nussinov R. 1999c. Distinguishing between sequential and nonsequentially folded proteins: Implications for folding and misfolding. *Protein Sci 8*:1591–1604.

Tsai CJ, Xu D, Nussinov R. 1998. Protein folding via binding, and vice versa. *Fold Des 3*:R71–R80.

Vonrhein C, Schlauderer GJ, Schulz GE. 1995. Movie of the structural changes during a catalytic cycle of nucleoside monophosphate kinases. *Structure 3*:483–490.

Wedemayer GJ, Patten PA, Wang LH, Schultz PG, Stevens RC. 1997. Structural insights into the evolution of an antibody combining site. *Science 276*:1665–1669.

Williams DC, Benjamin DC, Poljak RJ, Rule GS. 1996. Global changes in amide hydrogen exchange rates for a protein antigen in complex with three different antibodies. *J Mol Biol 257*:866–876.

Wolynes PG, Onuchic JN, Thirumalai D. 1995. Navigating the folding routes. *Science 267*:1619–1620.