1085-9195/99/31/141–164/$16.00

Folding Funnels and Conformational Transitions Via Hinge-Bending Motions

Sandeep Kumar,[1] Buyong Ma,[2] Chung-Jung Tsai,[2] Haim Wolfson,[3] and Ruth Nussinov[,1,4]*

[1]Intramural Research Support Program–SAIC, Laboratory of Experimental and Computational Biology, NCI-FCRDC, Frederick, MD, 21702; [2]Laboratory of Experimental and Computational Biology, NCI-FCRDC, Frederick, MD 21702; [3]Computer Science School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel; and [4]Sackler Institute of Molecular Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

ABSTRACT

In this article we focus on presenting a broad range of examples illustrating low-energy transitions via hinge-bending motions. The examples are divided according to the type of hinge-bending involved; namely, motions involving fragments of the protein chains, hinge-bending motions involving protein domains, and hinge-bending motions between the covalently unconnected subunits. We further make a distinction between allosterically and nonallosterically regulated proteins. These transitions are discussed within the general framework of folding and binding funnels. We propose that the conformers manifesting such swiveling motions are not the outcome of "induced fit" binding mechanism; instead, molecules exist in an ensemble of conformations that are in equilibrium in solution. These ensembles, which populate the bottoms of the funnels, *a priori* contain both the "open" and the "closed" conformational isomers. Furthermore, we argue that there are no fundamental differences among the physical principles behind the folding and binding funnels. Hence, there is no basic difference between funnels depicting ensembles of conformers of single molecules with fragment, or domain motions, as compared to subunits in multimeric quaternary structures, also showing such conformational transitions. The difference relates only to the size and complexity of the system. The larger the system, the more complex its corresponding fused funnel(s). In particular, funnels associated with allosterically regulated proteins are expected to be more complicated, because allostery is frequently involved with movements between subunits, and consequently is often observed in multichain and multimolecular complexes.

This review centers on the critical role played by flexibility and conformational fluctuations in enzyme activity. Internal motions that extend over different time scales and with different amplitudes are known to be essential for the catalytic cycle. The conformational change observed in enzyme-substrate complexes as compared to the unbound enzyme state, and in particular the hinge-bending motions observed in enzymes with two domains, have a substantial effect on the enzymatic catalytic activity. The examples we review span the lipolytic enzymes that are particu-

*Author to whom all correspondence and reprint requests should be addressed. E-mail: ruthn@ncifcrf.gov

larly interesting, owing to their activation at the water-oil interface; an allosterically controlled dehydrogenase (lactate dehydrogenase); a DNA methyltransferase, with a covalently-bound intermediate; large-scale flexible loop motions in a glycolytic enzyme (TIM); domain motion in PGK, an enzyme which is essential in most cells, both for ATP generation in aerobes and for fermentation in anaerobes; adenylate kinase, showing large conformational changes, owing to their need to shield their catalytic centers from water; a calcium-binding protein (calmodulin), involved in a wide range of cellular calcium-dependent signaling; diphtheria toxin, whose large domain motion has been shown to yield "domain swapping;" the hexameric glutamate dehydrogenase, which has been studied both in a thermophile and in a mesophile; an allosteric enzyme, showing subunit motion between the R and the T states (aspartate transcarbamoylase), and the historically well-studied lac represoor. Nonallosteric subunit transitions are also addressed, with some examples (aspartate receptor and *Bam*HI endonuclease). Hence, using this enzyme-catalysis-centered discussion, we address energy funnel landscapes of large-scale conformational transitions, rather than the faster, quasi-harmonic, thermal fluctuations.

Index Entries: Hinge-bending; lock-and-key vs induced-fit; conformational ensembles; binding; folding.

INTRODUCTION

Any molecule, including proteins, functions through its intermolecular associations. Hence, over the years, binding mechanisms have been the focus of intense research. In general, the mechanisms of binding have been classified according to the extent and type of the motion that is involved. Some molecular associations have been considered rigid, whereas others have been viewed as flexible, adapting the molecular structure to fit—and optimally bind—the incoming molecular partner.

Around the turn of this century, Emil Fischer proposed the "lock and key" mechanism for protein binding *(1)*. This mechanism can be illustrated best by the enzyme substrate binding process. The enzyme active site was thought to be a rigid and sturdy lock that could have an exact fit with only one substrate (key). Thus, the specificity of enzyme catalysis was thought to be the result of matching a lock with the correct key. This deceptively simple process was accepted to be a general mechanism for enzyme substrate binding for more than half a century, until challenged by an alternative mechanism of "induced fit" proposed by Koshland in 1958 *(2)*. According to the induced fit theory, proteins need not be perfectly rigid locks. Rather, they could accommodate the incoming substrate by flexibly adapting their substrate binding site. To a large extent, these two binding modes—rigid and flexible—have been subsequently distinguished through comparisons of the structures corresponding to the free, unbound, protein molecule to the corresponding structure when complexed with its ligand. If the structures are similar, the binding mode has been considered as belonging to the rigid, lock and key type mechanism. On the other hand, if the structural comparison has illustrated a relatively large conformational change, the binding mode has been classified as belonging to the induced fit category. These views of molecular associations have since then been widely accepted.

Recently, we have questioned these views on the mechanisms of protein binding. We have shown that in light of the "new view" of protein folding *(3–13)*, one can visualize the presence of an ensemble of conformational isomers of a protein in equilibrium with one another around the bottom of the energy funnel. During the binding of a substrate, the conformational isomer whose binding pocket shape is most complementary to the substrate conformation will be selected for binding. The equilibrium will then shift toward the conformation of the bound protein. This view of protein binding is fundamental to both the lock and key and the induced fit mechanisms of protein substrate binding since it combines

ideas of "complementarity" and "flexibility," central to the Fischer and Koshland theories. Hence, the energy landscape depicted by the folding funnel of a rigid protein that binds via the lock-and-key mechanism is likely to have a smoother bottom with a deep, well-defined minimum, as compared to the folding funnel of a flexible protein which binds through the so-called induced-fit mechanism. The latter may be depicted by a rugged funnel bottom, with multiple minima separated by low barriers. These correspondingly portray its range of conformational isomers.

Folding and binding are similar, hierarchical processes, involving intra- and intermolecular recognition and association *(14)*. The single difference between them is the presence (or, absence) of chain connectivity. Consequently, the funnel shape becomes more complex as the intra- or intermolecular units bind to each other. Single funnels fuse with each other, progressively increasing the funnel complexity. Hence, a binding mechanism illustrated by structural units which belong to the same molecule, can similarly also be manifested by subunits of an oligomer. In the former case, the most favorable conformational isomer of the single molecule binds the ligand. On the other hand, in the second case the most favorable conformer of the oligomer would bind. In both cases, the equilibrium will subsequently shift in favor of the binding conformer, whether monomer or multimeric complex, further driving the reaction. Fig. 1 depicts three types of folding/binding cases. In Fig. 1A, two rigid proteins (or structural units) can form a rigid complex or a semi-flexible complex; In Fig. 1B, a rigid protein binding to a flexible protein can result in a rigid complex, a semi-flexible complex or a highly flexible complex; and in Fig. 1C, the binding of two flexible proteins can result in a semi flexible complex or in a highly flexible complex.

Conformational flexibility and structural fluctuations play an important role in enzyme activity. A large variety of internal motions extending over different time scales and with different amplitudes are involved in the catalytic cycle. The conformational changes upon substrate binding, and particularly the hinge-bending motions that occur in enzymes consisting of two domains, have a substantial effect on the catalytic activity of enzymes.

Below we first discuss our view of binding mechanisms within the framework of the "new view" of protein folding. We show how this simple concept replaces long-held notions of induced-fit vs lock and key type binding. In particular, we illustrate how this framework applies equally well to intra- and intermolecular binding mechanisms. We proceed to provide a broad range of examples substantiating and corroborating our view. Several proteins are discussed extensively in order to relate structure, funnel, and reactivity. Our examples focus mainly on the hinge-type motions, frequently observed in protein associations. In light of the similarity between folding and binding, our examples progressively depict such hinge-bending motions between molecular fragments; between domains (or, compact hydrophobic folding units) and between subunits, substantiating the analogy between the folding and binding processes. We further describe experimental and computational approaches to detect such hinge-bending motions.

THE NEW VIEW OF PROTEIN FOLDING AND BINDING MECHANISMS

A "new view" of protein folding has emerged recently. It is the outcome of both progress in experimental methodology and, in parallel, a new conceptual framework. The most important point about the "new view" of protein folding is that folding does not proceed starting from a single non-native conformation through a single pathway toward the formation of the native protein structure. Such a single pathway notion was at the heart of the "old view" of protein folding. That view has lead to the long puzzling dilemma posed by the Levinthal paradox; namely, how would the single folding conformation be able to search

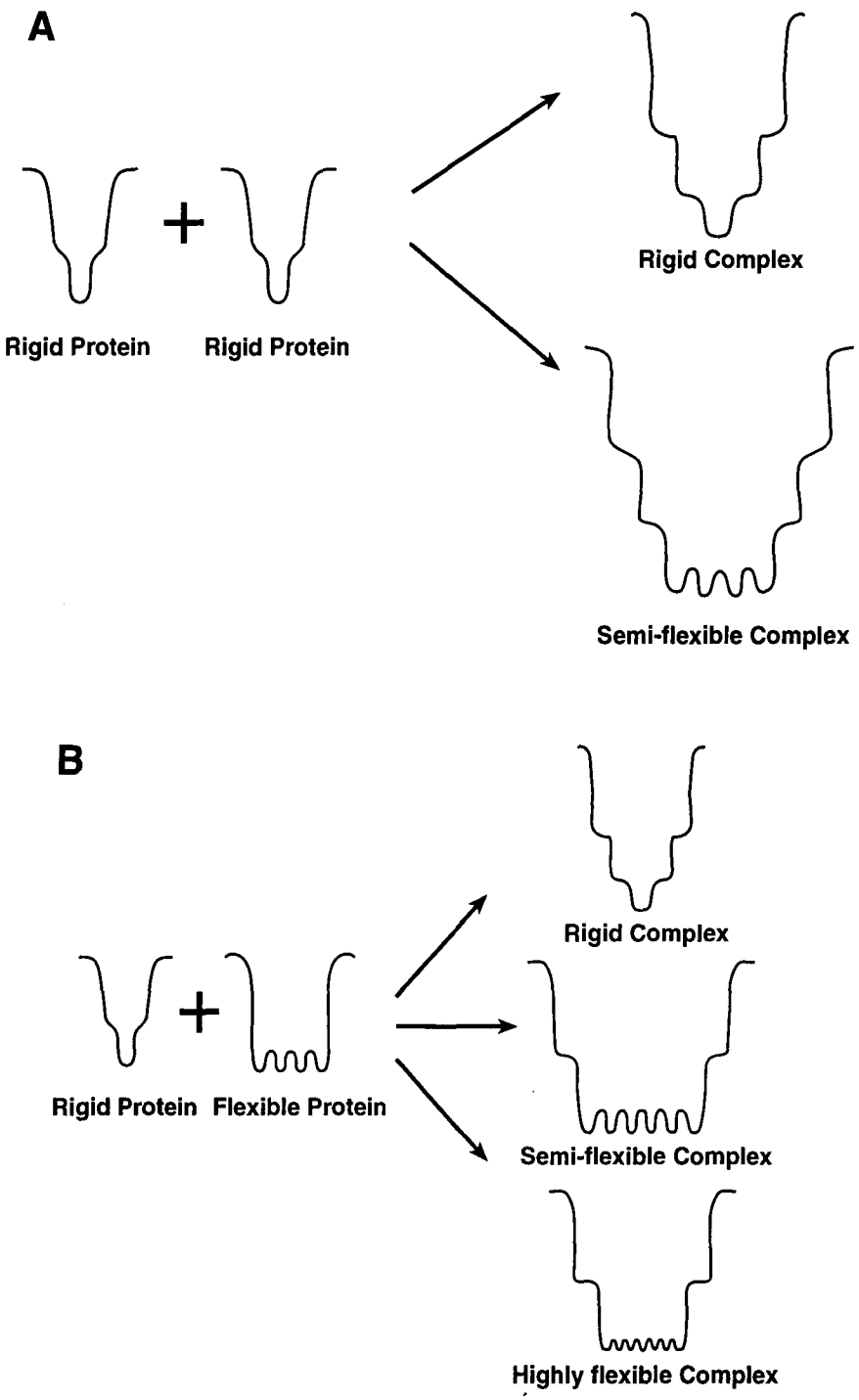
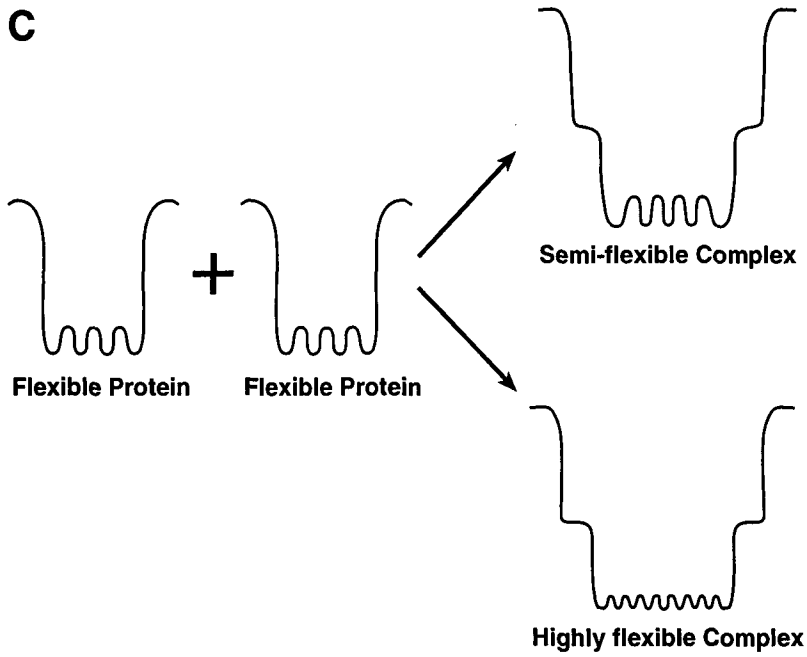


Fig. 1. The figure illustrates three types of folding/binding cases: (**A**) The binding of two rigid proteins can result in a rigid complex or in a semi-flexible complex. (**B**) The binding of a rigid protein and a flexible protein can result in a rigid complex, a semi-flexible complex or a highly flexible complex. (**C**) The binding of two flexible proteins can result in a semi-flexible complex or in a highly flexible complex.



through conformational space to reach its most stable structure within the allotted folding time span. In contrast, the new view has shown that this problem does not arise, as there are multiple starting conformations, proceeding through multiple pathways down the slopes of an energy funnel landscape, toward its bottom. This simple and attractive concept of multiple pathways gliding down the funnel landscape, supported by experimental evidence, recently has been discussed extensively in the literature. Much of the published discussion has centered on folding pathways, obligatory steps, transition states, and entropic barriers, seeking to capture the critical and elementary aspects of the protein folding problem.

Yet, as noted previously, proteins function through their intermolecular binding. And, although much attention has focused on the multiple pathways rolling downhill, the steepness of the slopes, and the bumpiness of the funnel walls, the implications for long-held notions on binding mechanisms have not been considered. Here our goal is to focus on the bottom of the funnels and their implications. If the proteins are rigid, one may envision that around the bottom there would be a single, or a few, minima. And, conversely, flexible molecules would display rugged funnel bottoms with low barriers, corresponding to a range of conformational isomers.

The binding mechanism is often assigned via a comparison of the structures crystallized both of the bound, complexed, conformer and of the unbound, free conformational isomer. If a relatively large difference between the two conformers is observed, the mechanism has generally been classified as an induced fit. On the other hand, if only a small difference is detected, the mechanism is considered to belong to the rigid, lock-and-key-type association. Yet, if the molecule exists in an ensemble of conformers, it is easy to imagine that the conformer that is complementary to the ligand would bind to it. Since the conformers are in equilibrium in solution, there would be a shift in favor of that conformer. In the case of the crystal of the 'unbound' conformer, we should bear in mind that this isomer is actually also bound, except that it is in association with its twin molecule. Hence, the conformer whose structure is determined here is the one complementary to its twin in the crystal. On the other hand, the conformer that is crystallized

in the so-called "bound," complexed case is the conformer that is complementary to its cognate ligand. Thus, a difference between the two conformations, which is observed while superimposing the bound and unbound structures does not necessarily mean that in binding to the ligand the protein has undergone an induced fit mechanism. Rather, it is the outcome of crystallizing different conformers. We stress, however, that here we discuss conformational changes observed in backbone movements. On the local level, one may easily imagine that side-chain optimization between the interacting molecules would take place, to achieve an optimized stable association.

Hence, the presence of an ensemble of conformational isomers around the bottom of the folding funnel rationalizes and generalizes binding mechanisms. Rather than resorting to lock-and-key, induced fit, and, similarly, to crystal effects, swapped dimers, and amyloid formation, we should simply consider the variability of the conformers. In all of these cases, the conformer that binds is the one most favorable and complementary, either to its ligand, to the growing crystal, or to the amyloid fiber, with the equilibrium shifting in its favor. Different conformers may well bind different ligands. Hence, the larger the flexibility, the broader the range of binding specificity. Furthermore, binding and folding are similar processes. Hence, this conceptual approach applies equally well to binding funnels. At the bottom of the binding funnel there may similarly be an ensemble of conformations of bound, multimolecular assemblies. These will further bind the incoming ligand. Here, the already bound, multimolecular conformer—whether composed of identical subunits or of different molecules bound in a complex—whose structure is most favorable for further binding, is the one that would associate. Hence, the principles are general, regardless of whether in the comparisons of the unbound and bound structures we observe movements of fragments of the chain, of secondary structure elements, of structural domains, or of subunits.

Below, we illustrate the applicability of our argument to one particular type of conformational transition. We further rationalize our arguments with respect to allosteric and motions that do not involve allosteric transition.

HINGE-BENDING MOTIONS

Protein motions have been very conveniently systematized on the basis of packing *(15,16)*. Packing has been selected as the basic criterion since atoms are very tightly packed in the interior of the protein molecule. Groups of atoms can move with respect to each other only if there is a packing defect, or a cavity, that enables them to do so. Interfaces between groups of atoms or between structural parts are not smooth. Interdigitation, particularly of side-chains, imposes severe constraints on the movements of the structural units if their internal packing is to be preserved. In hinge-bending motions, the structural units move with respect to each other; however, while the packed arrangement within the unit is conserved, the packing at their interface is disrupted. The parts move as relatively rigid bodies with respect to each other, swiveling on their "hinge." The motion observed is roughly perpendicular to the interface.

Hinge-bending movements are obliquely opposite to those classified as "shear." In shear movements, the packing at the interface is maintained with the structural units sliding with respect to each other. Additionally, whereas in hinge-bending movements a few large (twisting) changes may be observed, in shear movements many small changes parallel to the plane of the interface may be seen *(15)*. Here we confine ourselves largely to hinge-bending motions. Such movements may be observed between small structural parts, such as loops, or secondary structure elements, between hydrophobic folding units, between domains, or between subunits. Hence, they can be intramolecular, connected by the backbone of the polypeptide chain, or they can be intermolecular. Thus, in this sense too, intra- and intermolecular associations resemble each

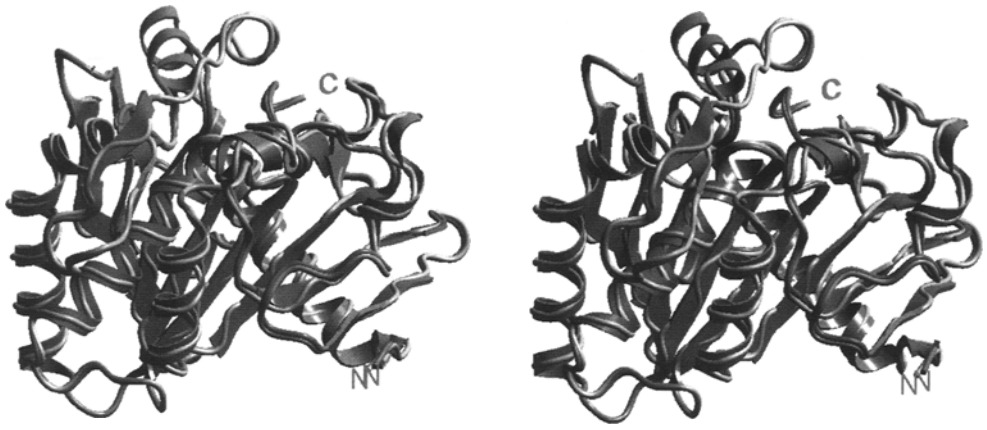


Fig. 2. An example of a superpositioning of a fragment–hinge-based motion. Shown here, the triglyceride lipase 1tgl vs 4tgl in PDB *(19)*.

other. The only exception is that owing to the chain connectivity in intramolecular hinge motions, there are additional constraints on the movements imposed by the covalently linked polypeptide chain.

SPECIFIC EXAMPLES

Here we provide examples of hinge-bending type motions. The examples are given for the three major classes: motions of fragments, motions of domains, and motions of subunits. We describe the hinge bending involved (sometimes with an accompanying illustration) superimposing the open and closed conformers.

Fragment Motion

TRIGLYCERIDE LIPASE

Lipolytic enzymes are particularly interesting, due to their activation at the oil–water interface. Various hypotheses have been proposed to explain their interfacial activation. These lipases are serine hydrolases, composed largely of mixed β-sheets, with the strands predominantly in a parallel orientation. The catalytic triad of the fungus lipase is composed of the Ser-His-Asp residues, with the nucleophilic Ser located in a sharp turn between one of the β-strands and a buried α-helix *(17)*. The active center is buried beneath a short amphipathic helix. This has led to the suggestion *(18)* that triglyceride lipases may undergo a conformational change in response to adsorption at the oil–water interface. Detailed analysis of the stereochemical conformational change that takes place during enzyme activation under these conditions has shown a movement of a 15-residue-long "lid" in a hinge-type rigid body motion. This results in transporting the short α-helix by over 12Å. This movement is the outcome of two hinge regions on both sides of the α-helix. In particular, this swiveling exposes a large (800Å^2) hydrophobic surface. In the process, one residue changes its conformation from extended to an α-helix. That the fragment moves as a rigid body is illustrated by the low RMSD which has been measured for these residues between the two (active and inactive) conformations (0.31Å). Hence, this is a typical rigid-body hinge bending motion of a fragment. These two conformations (PDB codes 1tgl and 4tgl; ref. *19*) are superimposed in Fig. 2. The conformation of the lid has been suggested to be stabilized by a new set of hydrogen bonds,

with a plethora of possible meta-stable conformers *(17)*. This protein can be visualized as having relatively simple folding and binding funnels, with only a few minima separated by low energy barriers around the bottom. The minima at the bottom will have similar free energies. Such cases correspond to the schematic representation in Fig. 1A.

The activation pathway of three wild-type lipases and three mutants were investigated using molecular dynamics techniques *(20)*. The results suggest that the activation of lipases is enhanced in a hydrophobic environment, consistent with experimental observations. The energy gain upon activation varies between the different lipases, and depends strongly on the distribution of the charged residues in the activating loop region. In a low dielectric constant medium (such as a lipid environment), the electrostatic interactions between the residues located in the vicinity of the activating loop (lipid contact zone) are dominant and determine the activation of the lipases. Some titratable residues experience significant pK shifts upon activation. In general, the active site has to be open for function *(21)*. The colipase appears to stabilize the lid domain of the lipase when it is in the open conformation, and thereby facilitate lipolysis *(22)*.

Lactate Dehydrogenase

The crystal structures of the T and L states of lactate dehydrogenase from the bacteria *B. longum*, have illuminated a mechanism for allosteric control *(23)*. When the enzyme is bound to lactate and nicotinamide adenine dinucleotide (NAD), a large conformational change is observed. This change entails a movement of a surface loop by roughly 10Å, covering the active site. Additionally, five helices and three loops are observed to alter their position, moving by about 2Å. The loop moves on two hinges. The first illustrates large changes in two torsion angles, in a typical hinge motion. The second hinge, however—which is located in a helix—is constrained. To accommodate this rotation, the deformation is distributed over several torsion angles, and splits the α-helix into two helical components, with different side-chain packing *(16)*. The moving parts are on the surface of the protein, with the motion of the loop propagating away from the substrate binding location, with increasing mobility. This is a classic example of a highly flexible protein adopting different conformations in the T and L states. The folding and binding funnels for lactate dehydrogenase are considerably more complex than those for triglyceride lipase. These funnels contain a large number of minima separated by low energy barriers at their bottoms. These minima allow the observed large conformational changes in the bound and unbound states. The funnels for lactate dehydrogenase can be schematically represented by Fig. 1C.

HhaI Methyltransferase

The crystal structure of the methyltransferase covalently bound to an intermediate has been determined. The DNA is located in a cleft between two domains of the enzyme. It has a characteristic B-form conformation, with the exception of the disrupted G-C base pair, which contains the target cytosine. The cytosine is observed to have swung out of the helix. The enzyme also illustrates a conformational change: two hinges enable a 20-residue active-site loop movement, resulting in a 25Å shift. This loop, containing six highly conserved residues, moves toward the DNA, making contacts with the DNA via its minor groove. The interaction is with the backbone of the DNA. Interestingly, while this interaction and movement is in the large domain, the small domain illustrates a small, concerted movement toward the binding cleft of the DNA *(24)*. This indicates that the enzyme and substrate are semi-flexible. The energy funnels are likely to be simpler than those in the case of lactate dehydrogenase, but more complicated than those of triglyceride lipase. Since the DNA retains its B-form (except for one base pair), it can be thought of as being relatively rigid. On the other hand, Hha1 methyltransferase can be thought to contain flexible and rigid domains. The energy funnels for this case can be schematically represented by Fig. 1B.

TRIOSE PHOSPHATE ISOMERASE (TIM)

TIM is a glycolytic enzyme which catalyzes the interconversion of glyceraldehyde-3-phosphate and dihydroxyacetone phosphate. This enzyme has been well studied, and the open and closed forms of the TIM loop have been analyzed in some detail. The loop has two hinges. The C_α's have been shown to move by about 7Å. TIM is a β-barrel enzyme, with eight helices arranged around a barrel of eight strands. The 11-residue loop connects an inner strand with an outer helix. This flexible loop is observed to be closed when the substrate is bound at the active site of the enzyme. In particular, the hinge of the loop has been studied by extensive mutagenesis. The sequences at the hinges in these mutants have revealed that the solutions to the hinge flexibility problem are quite varied, with the preferences being sequence-dependent. In particular, Gly, the smallest residue lacking a side chain, is avoided, suggesting that unrestricted hinging is selected against if the enzyme is to be biologically functional *(25)*. TIM illustrates an example of a protein having a relatively smooth energy landscape with a few minima around the bottom. The enzyme retains its overall geometry in the unbound and complexed forms, except for small loop movements. It conforms to the schematic illustration of Fig. 1A.

TIM usually functions as a dimer. The monomeric form of TIM is inactive and thermodynamically unstable *(26)*. The unfolded TIM monomers are susceptible to proteolytic digestion and thiol oxidation, while native TIM is resistant to both. The dimeric nature of TIM decreases the frequency of subunit unfolding by several orders of magnitude, thereby increasing the chemical stability of the protein *(27)*. Stability appears to be the main reason for the dimerization. However, a stable monomeric variant of trypanosomal TIM (monoTIM) has also been reported. MonoTIM has lower, but significant, catalytic activity *(28)*. In the dimer, the two subunits function independently, and the conformational change is largely intra-subunit *(29)*. Hence, the funnel of TIM is the result of the coupling of the funnels of its two subunits. The coupling, even though weak, is crucial for the dimer's stability.

To explore the origin of the large-scale motions of the flexible loop in triosephosphate isomerase (residues 166 to 176) at the active site, several simulation protocols have been employed both for the free enzyme *in vacuo* and for the free enzyme with some solvent modeling. These protocols include high-temperature Langevin dynamics simulations, sampling by a "dynamics driver" approach, and potential-energy surface calculations. The simulations and analyses indicate that in the context of a spontaneous opening of the free enzyme, the motion is a rigid-body type. The specific interaction between residues Ala176 and Tyr208 was found to play a crucial role in the loop opening/closing mechanism *(30)*.

Domain Motion

PHOSPHOGLYCERATE KINASE

3-Phosphoglycerate kinase (PGK) is essential for most living cells, both for ATP generation in the glycolytic pathway of aerobes and for fermentation in anaerobes. In addition, in many plants the enzyme is involved in carbon fixation. Like other kinases, PGK folds into two distinct domains, which undergo a large hinge-bending motion upon catalysis, with an active site located in the interdomain cleft. The monomeric enzyme catalyzes the transfer of the C1-phosphoryl group from 1,3-bisphosphoglycerate to ADP to form 3-phosphoglycerate (3-PG) and adenosine triphosphate (ATP). For decades, the conformation of the enzyme during catalysis has been enigmatic. The large distance between the binding sites for 3-PG and ATP, deduced from the crystallographic structures of the binary complexes, gave rise to the hypothesis that this enzyme undergoes a hinge-bending domain motion from an open to a closed conformation during catalysis.

However, direct experimental evidence for the closed conformation in the presence of both substrates has appeared only recently *(31,32)*. The crystal structure of PGK from the hyperthermophilic *Thermotoga maritima* (TmPGK)

represents the first structure of an extremely thermostable PGK *(31)*. The crystal structure of TmPGK was determined to 2.0Å resolution, as a ternary complex with two products, 3-phosphoglycerate and the adenylyl-imido diphosphate (AMP-PNP) analogue. The complex crystallizes in a closed conformation with a drastically reduced inter-domain angle. The distance between the two bound ligands is 4.4Å. The closed form represents the active conformation of the enzyme. The structure provides new details about the catalytic mechanism of TmPGK. An inter-domain salt bridge between residues Arg62 and Asp200 forms a strap that holds the two domains in the closed state. Lys197 was identified as a residue that is involved in stabilizing the transition state phosphoryl group. Comparison of TmPGK with its less thermostable homologues, reveals that the higher rigidity of TmPGK, is achieved through a larger number of intramolecular interactions, such as an increased number of ion pairs and further stabilization of the helix–loop regions *(31)*. The crystal structure of an additional ternary complex of PGK has been solved recently *(32)*. In this structure, the PGK ternary complex exhibits a dramatic closing of the large cleft between the two domains. This closure brings the two ligands, 3-phosphoglycerate and ADP, into close proximity. The availability of this structure has confirmed that PGK is a hinge-bending enzyme.

PGK represents a protein with strong domain–domain interaction. Cheung and Mas *(33)* have studied several PGK mutants with single tryptophans placed at various locations. These were used as intrinsic fluorescent probes to examine the extent and delocalization of conformational changes induced upon the binding of 3-PG, 1,3-diPG, ADP, ATP, and PNP-AMP (nonhydrolyzable analogue of ATP), and upon the simultaneous binding of 3-PG and PNP-AMP. The results showed that only those probes situated in the hinge, and in the parts of each domain close to the hinge, reflect substrate-induced conformational changes. The binding of substrates to one domain was found to induce spectral perturbation of the probes in the opposite domain, indicating a transmission of conformational changes between the domains. A combination of both substrates generated significantly larger fluorescence changes than the individual substrates.

The strong domain–domain interaction is also observed in the activity of PGK. In order to determine the role of the C-terminal helix in the folding and stability of yeast phosphoglycerate kinase, a mutant wherein the 12 C-terminal residues were deleted (PGKD404–415) has been constructed. This mutant folds into a conformation very similar to that of the wild-type protein, yet it exhibits a very low activity. The main structural effect of the deletion of the C-terminal helix is an increase in the flexibility of the whole protein, and a concomitant decrease in stability. The structural properties of the truncated protein are very similar, at least qualitatively, to those of the isolated domains. The C-terminal part of yeast PGK kinase is not necessary for most of the initial folding steps. However, it acts to lock the C-domain onto the N-domain, thus ensuring full-enzyme activity *(34,35)*. Nevertheless, the domain–domain interactions are best described as the sensitivity of the PGK activity to conformational changes *(36,37)*, though not in terms of domain–domain interaction energy. The actual interdomain contacts are relatively weak. Interactions relating to folding pathways are those within domain cores. Contacts formed either between domains or with the interdomain helix are made only in the folded ground state. They do not constitute a separate step in the folding mechanism *(38)*. A large number of studies are in agreement that the folding-unfolding of PGK is a sequential, multistep process *(36–39)*. It is particularly interesting to note that the hinge motion already exists in the folding intermediate *(40)*. Thus, the funnel of PGK illustrates a simple merging of the two corresponding sub-funnels of its domains.

Despite overwhelming evidence regarding the hinge bending in PGK, McPhillips et al. *(41)* have found a different case. The structure of a ternary complex of the R65Q mutant of yeast PGK with magnesium 5′-adenylylimidodiphosphate (Mg AMP-PNP) and 3-PG has been determined by X-ray crystallography

to 2.4Å resolution. Yeast PGK also consists of two domains, with the 3-PG bound to a patch of basic residues from the N-terminal domain, and the Mg-AMP-PNP interacting with residues from the C-terminal domain. The two ligands are separated by about 11Å across the interdomain cleft. The model of the R65Q mutant of yeast PGK is very similar to the structures of PGK isolated from horse, pig, and *Bacillus stearothermophilus*. The most significant tertiary structural differences among the yeast R65Q, equine, porcine, and *B. stearothermophilus* PGK structures occur in the relative orientations of the two domains. However, the relationships between the observed conformations of Yeast PGK are inconsistent with a hinge-bending behavior that would close the interdomain cleft. Hence, it has been proposed that the available structural and biochemical data on yeast PGK may indicate that the basic patch primarily represents the site of anion activation and not the catalytically active binding site for 3-PG *(41)*.

Adenylate Kinase

Kinases have to shield their catalytic centers from water, to avoid becoming ATPases. Hence, large differences are observed between the bound and the unbound structures *(42)*. There are two interdomain linkages and four hinges. The first pair of hinges yields a 60° rotation, and the second pair shows a rotation of 30°. Hence, the total rotation of the domain is around 90°. This mobile domain is connected to the rest of the molecule by two helices, oriented in an antiparallel fashion, and packed against each other. In the closed form, deformations near the hinges are observed, at both termini of the helices. These involve rotations of three torsion angles at each joint. The regions between the hinge-joints, the two helices and the remainder of the mobile domain move as relatively rigid bodies. Energy landscapes and funnels for proteins exhibiting domain motion can, in general, be expected to have higher degree of complexity. Adenylate kinase exhibits a funnel that can be thought to arise out of a fusion of two energy funnels. One of the funnels is smooth with relatively few minima around the bottom. This corresponds to the rigid domain of adenylate kinase. The mobile domain has a complex energy funnel with several isoenergetic minima around the bottom of the funnel. In this sense, funnels for adenylate kinase can be schematically represented by Fig. 1B. The fusion of these two funnels gives rise to a semi-flexible adenylate kinase funnel, capable of producing motion in the mobile domain, independent of the rigid domain. The observed conformational change can then be rationalized on the basis of such a funnel.

Adenylate kinases undergo large conformational changes during their catalytic cycle. Domain motions occur largely independently from each other *(43)*. Muller et al. *(44)* have solved the structure of the unligated adenylate kinase and compared it with the structure of the same enzyme, ligated to an inhibitor mimicking both substrates, ATP and AMP. This comparison has shown that, upon substrate binding, the enzyme increases its chain mobility in a region remote from the active center. Zhang et al. *(45)* have also found that the change of the enzyme activity coincided with that of the rate of ANS binding during denaturation by low concentration of the denaturants, suggesting that the activation of adenylate kinase by denaturants may be due to an increase in conformational flexibility at its active site.

Calmodulin

Calmodulin is a ubiquitous Ca^{2+} binding protein. It is involved in a broad range of cellular Ca^{2+}-dependent signaling processes. It regulates the activity of a large number of proteins, such as the kinases, phosphatases, nitric oxide synthase, inositol triphosphate kinase, nicotinamide adenine dinucleotide kinase, cyclic nucleotide phosphodiesterase, Ca^{2+} pumps, and proteins active in motility *(46)*. Comparison of the complexed (with M13 peptide) nuclear magnetic resonance (NMR) structure and an unbound crystal structure has shown the existence of a hinge-based motion. This motion involves the splitting of a long helix into two helices, kinked with an approx

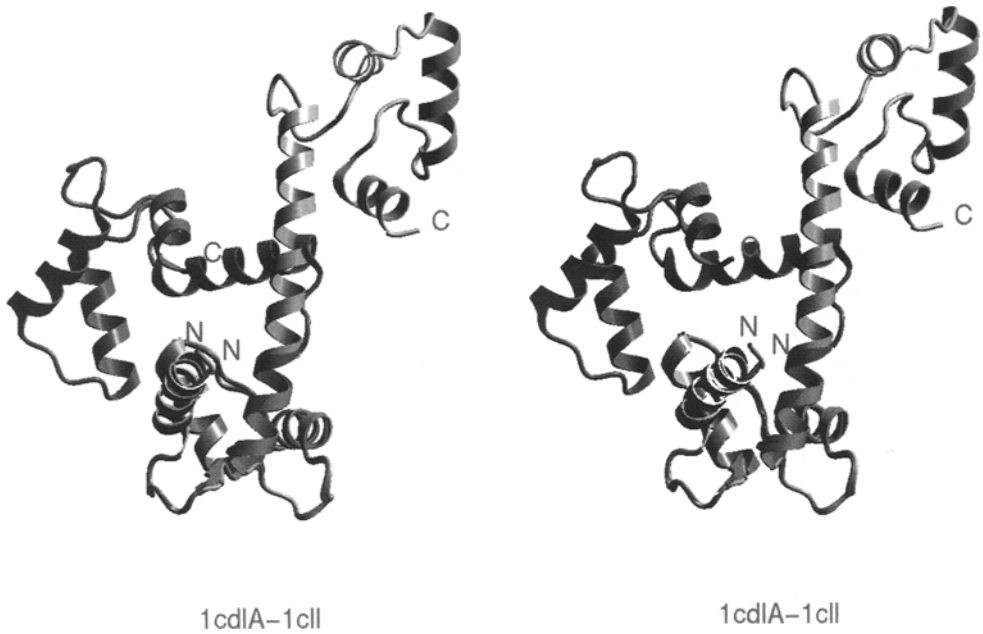


Fig. 3. An example of a superpositioning of a domain–hinge-based motion. Shown here, the calmodulin, PDB codes 1cdl (chain A) and 1cll.

100° angle in between. The bound structure shows that the two domains closing around the peptide ligand also make contacts with each other. The long interdomain helix is only marginally stable in solution. In breaking into two component helices, it partly unfolds around the four-residue hinge, which adopts an extended conformation. The additional twist around the helix axis results in a total rotation of the domains of approx 150° with respect to each other. Changes in the hinge-joint region, and side-chain movements, allow binding variability. Hence, this is an example of flexibility enabling the binding of a range of peptides with different sequences. Fig. 3 depicts this hinge-bending domain movement (PDB codes 1cdl and 1cll). Calmodulin can be expected to have a complex energy funnel with a large number of isoenergetic minima around funnel bottom, allowing large movements. These funnels can be represented by Fig. 1C.

Diphtheria Toxin

Diphtheria toxin is a protein secreted from the bacteria that cause diphtheria. Diphtheria toxin is active as a monomer. It does not convert to a dimer even at high concentrations. However, changes in conditions (temperature and pH) cause a globular domain to swing out of the body of the protein. The open monomer subsequently dimerizes via "domain swapping" with a second monomer. There is, then, one interdomain linkage, and a single hinge. The domain rotates about the hinge resulting in 180° rotation with respect to the closed structure. Comparison of the two structures illustrates a 65Å movement *(47)*. The energy funnels for diphtheria toxin can be represented by Fig. 1B. The funnels are expected to be much more complicated than those of adenylate kinase and annexins. The large structural change observed between the open and closed

forms can be explained by the presence of several minima around the bottom of the funnel that are similar in energy but can describe geometrically distinct conformations.

Glutamate Dehydrogenase

Glutamate dehydrogenase catalyzes the reversible oxidative deamination of L-glutamate to 2-oxoglutamate and ammonia, with NAD^+ or $NADP^+$ as a cofactor. Each subunit of this hexameric enzyme contains two domains, which are separated by a deep cleft. One domain has been proposed to direct the self-assembly of the molecule into a hexameric oligomer *(48)*, whereas the structure of the other resembles the classical dinucleotide binding fold; but with one strand reversing its direction. The complexed structure shows that the nicotinamide portion of NAD^+ is bound deep in the cleft between the domains. A small conformational change is required to orient this moiety close to the substrate for catalysis to take place. A 13° rotation of one domain with respect to the other, via a hinge-bending type motion achieves this goal. Recently, a comparison of Glutamate dehydrogenases from *Pyrococcus furiosus*, a hyperthermophilic archaebacterium, and meso-philic *Clostridium symbiosum* has shown that the thermophilic glutamate dehydrogenase has several additional salt bridges around the active site in each subunit *(49)*. The energy funnel for each subunit of this enzyme is again expected to be the result of fusion between the funnels for the two domains as in the case of adenylate kinase. Thus, the funnel of each subunit can be schematically represented by Fig. 1B. However, the energy landscape and funnels for the whole enzyme are expected to be highly complex, arising from the fusion of each of the funnels of the subunits.

GroEL Domain

GroEL belongs to the class of the chaperonins. These are large, multisubunit assemblies, critically important in mediating ATP-dependent protein folding. The GroEL belongs to the Hsp60 class. Members of this class are found in the bacterial cytoplasm (like GroEL), in the mitochondria (Hsp60) and in the chloroplast (like rubisco-binding protein). Like other members of this family, GroEL consists of two stacked rings, with each ring containing seven radially-arranged subunits *(50)*. GroEL shows both domain and subunit hinge-bending movement. Here we discuss the domain motion.

There are three domains in each subunit. The intermediate domain swings down by approx 25° toward the equatorial domain and the central channel. The apical domain swivels up by approx 60° relative to the equator, and rotates around the axis of the domain by about 90°. This results in the interaction of the mobile loop of the GroES cap. This movement is around a pair of Gly (192 and 375) between the intermediate and the apical domain. The movement couples the binding of the GroES to the presence of the ATP. Both intermediate and apical domain movements are of the rigid-body hinge-bending type. Comparison of the bound and unbound conformations shows RMSD's of 0.91Å and 1.66Å, respectively, for the intermediate and apical domains.

Subunit Motion

Aspartate Transcarbamoylase

Aspartate transcarbamoylase is an allosteric enzyme, and its motion is an example of allosteric transition *(51)* between R (relaxed) and T (taut) states. The N-terminal region of the regulatory subunit is important for controlling nucleotide binding, creating the high-affinity and low-affinity effector binding sites, and coupling the binding sites within the regulatory dimer *(52)*. The enzyme catalyzes the reaction between carbamoyl phosphate and L-aspartate to yield phosphate and N-carbamoyl-aspartate. The latter is a precursor in the synthesis of pyrimidines. Aspartate trans-carbamoylase consists of six regulatory subunits, and of six catalytic subunits. The latter, catalytic subunits, are arranged as a dimer of two trimers. On the other hand, the six regulatory subunits are arranged as three dimers, each connecting the two sides of the catalytic trimer. Movement from a T to an R state involves rotation of approx 15° of the regulatory

dimer subunits, around their dimer axis. This movement is accompanied by a 12Å separation of the catalytic trimer, and a 10° rotation about the three-fold axis. Each catalytic subunit is composed of two domains, whose movement is coupled to the overall subunit motion. Hence, overall, this is an example of concerted subunit-domain motions. The energy landscapes in both the folding and binding funnels, which correspond to proteins showing subunit motions, have a much higher degree of complexity than those of proteins showing fragment and domain motions. Yet, owing to the analogy in the precesses of folding and binding, these funnels can be schematically illustrated in a simple form. Hence, for example, the energy funnels of aspartate transcarbamoylase can be represented by Fig. 1A. Each subunit is rigid and has a simple energy landscape; however, the complex is semiflexible which allows subunit motion between the T and R states of the enzyme.

Lac Repressor

The lac operon, and the allosteric regulation involved, has been a paradigm of gene regulation for many years. The lac repressor has been crystallized both bound to an inducer (isopropyl-β-D-1-thiogalactoside) and complexed with a 21-base pair symmetric operator DNA, in addition to the free, unbound state *(53)*. Comparisons of the three structures show that there are two distinct structural rearrangements of the monomer and the dimer. These correspond to the induced and the repressed states. The conformation of the monomer is similar to that observed when complexed with the inducer. However, comparison of the free, unbound structure of the repressor with its conformation when complexed with the DNA operator shows that the repressor has undergone a change. A hinge-like motion is observed within the dimer, altering the interface between the NH_2 terminal subdomains. The two subdomains in the dimer are rotated by about 10°, and their centers of mass move by about 2Å. This hinge motion does not alter the relationship between the two COOH terminal subdomains. Fig. 4 depicts this structural change (PDB codes 1lbi and 1lbg). Lac repressor can be represented by Fig. 1B. The subunits are flexible and thus yield a flexible complex.

Non-Allosteric Transitions

While the subunit hinge-based motions previously described are all related to allosteric transitions, there are other, similar hinge-bending movements between subunits that are not linked to allostery. The following are some examples.

Aspartate Receptor

The aspartate receptor of chemotaxis belongs to a large class of transmembrane proteins. These proteins contain an extracellular, ligand-binding domain, a cytoplasmic signaling domain, and a transmembrane domain. The ligand-binding domain has been crystallized in the presence and in the absence of aspartate *(54)*. There are two subunits. Each subunit is a 4 α-helix bundle, with two long NH_2 and COOH terminal helices, and two shorter helices. These form a cylinder of 20Å in diameter, and 70Å long. The substrate binding site is at the interface between the two subunits. Comparison between the free and the bound structures has indicated that although the structures of the subunits are largely unchanged (except for some change in the conformation of one loop), the subunits change their orientation with respect to each other. The conformational change is apparently propagated through the membrane, and is involved in signal transduction.

*Bam*HI Endonuclease

*Bam*HI endonuclease is a restriction enzyme, that cuts the DNA at a specific sequence. There are two subunits in the enzyme structure. A number of differences have been observed between the bound and the unbound conformations *(55)*. In particular, the subunits rotate with respect to each other by about 19°. In addition, there are some local changes involving 12 residues between two strands, and a

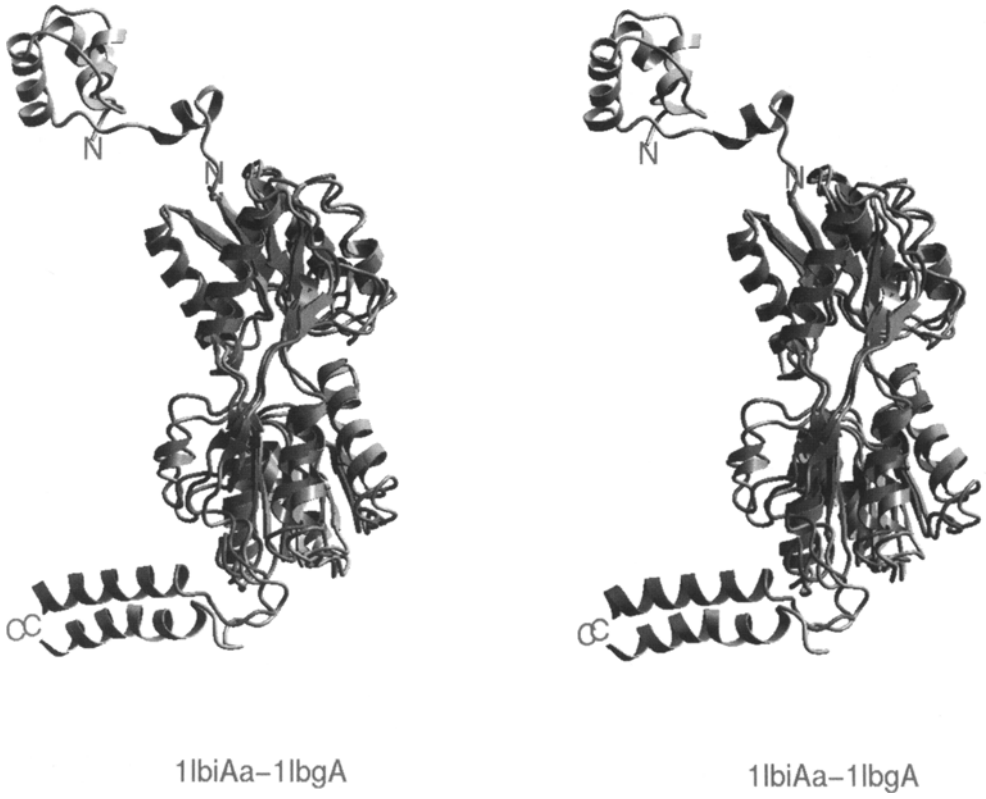


Fig. 4. An example of a superpositioning of a subunit–hinge-based motion. Shown here, the lac repressor PDB codes 1lbi and 1lbg.

rearrangement of a 5-residue loop. Some unfolding of the C-terminus helix into an extended structure is also observed.

Additional changes between subunits that do not involve allosteric transitions are observed, for example in the immunoglobulin VL-VH movements *(56–58)* and in *Saccharomyces cerevisiae* PPR1 Zn-finger DNA recognition protein *(59)*.

The examples given above describe a wide range of hinge-bending motions observed in proteins at different structural levels. Despite the varying complexity of the energy funnels for the proteins in these examples, it is clear that most of these can be described in the simple schematic terms outlined in Fig. 1. This indicates the similar hierarchichal nature of protein-folding and protein-binding processes. Furthermore, viewed this way, we are able to understand and to rationalize protein function based on the simple concept of energy landscapes and the new view of protein folding.

MEASUREMENT AND ANALYSIS OF CONFORMATIONAL TRANSITIONS VIA HINGE-BENDING MOTIONS

Hinge-bending motion is a large amplitude collective motion. Low-frequency normal modes may become useful for determining a first approximation of the conformational path between the closed and open forms of these proteins *(60)*. The low-frequency normal modes could be as low as 4.1 cm^{-1} for mouse epidermal growth factor (mEGF) *(61)*. This

hinge-bending motion corresponds to the "mitten mode." In this motion, the N-terminal domain is almost rigid. However, the C-terminal domain is found to consist of three rigid segments. Two segments are observed to move in the same direction; However, the third segment moves in the opposite direction. For this mode, the effective Young modulus is 1.1×10^9 dyn.cm^{-2}. This value is a little larger than that of the mode with the lowest frequency, 4.4 cm^{-1}. The difference may be related to the fraction of residues involved in β-sheets in the molecule. Similarly, a normal mode analysis found that the hinge-bending motion of the PGK has a frequency below 5 cm^{-1} *(62)*. These authors also found three types of movements: a propeller twist motion, a scissors-type hinge motion, and a shear motion between the domains. Local conformational variations may couple to the domain motions.

The experimental method that usually detects hinge motions is X-ray crystallography. Normal mode analyses or molecular dynamics simulations can also be useful in detecting molecular motions *(63,64)*. Recently, additional experimental methods have been utilized in studies of hinge bending. One of these is an application of the "essential dynamics" method to the cluster of structures obtained by NMR *(65)*. Using infrared and radio frequency to detect low frequency motion has also been proposed. For the purpose of the experimental measurements of the space dependence of the absorption, the direction of the incident wave and its polarization are calculated with respect to the main crystal planes in the case of maximal efficiency of the absorption. Anisotropy of the absorption allows both to distinguish exactly the normal mode causing this absorption, and to obtain some unknown parameters of the mechanical motion from the experimental data *(66)*.

A dielectric relaxation peak due to an intramolecular motion in the active site of trypsin was first observed in aqueous solution below the freezing temperature of bulk water by a time domain reflectometry method *(67)*. If trypsin inhibitor is added to the solution, it vanishes. It has been suggested that the motion observed is a hinge-bending deformation, giving rise to the enzymatic activity of trypsin. It is prohibited by a linkage of the trypsin inhibitor.

DETECTION AND PREDICTION OF CONFORMATIONAL TRANSITIONS VIA HINGE-BENDING MOTIONS

Is there any way to *a priori* pinpoint the region which can undergo such a conformational transition? If the two structures, i.e., the open and the closed conformations, are known, as in the examples shown previously, then one can straightforwardly superimpose them. If there is a single hinge, such a rigid-body superposition would yield two potential matches, as either the first parts of the two conformations, or the second substructural parts, would superimpose. Combining the two matches yields the position of the hinge. If, however, only one of the structures is available, then predicting the hinge location is a much more difficult task. One strategy is to inspect a set of structures belonging to the same family, or superfamily, carrying out structural comparisons, seeking similarity between family members. In such comparisons, the substructures may be rotated with respect to each other, yielding clues to the location of the hinge. However, here the extent of similarity between the structures may be variable, and the way to superimpose the structural pairs may be unclear.

Chandra et al. *(68)* attempt to allow far natural directional bias within the closed protein by repeatedly applying a weak pulling force over a short distance between pairs of atoms chosen at random in the two domains in question. Appropriate parameters governing the pulling function are determined empirically. The method was applied to the bilobal protein PGK and a closed-form activated ternary complex, generated for the *B. stearothermophilus* PGK. The model predicts the correct hinge regions, although the magnitude of the movement at one of the hinge points was overestimated. This model provides a reasonable representation of the closed-form ternary complex.

Methods developed originally to analyze domain motions from simulation *(63)* have been adapted and extended for the analysis of X-ray conformers of proteins with more than two domains. The methods can be applied as an automated procedure to any case where more than one conformation is available. The basis of the methodology is that domains can be recognized given the difference in the parameters governing their quasi-rigid body motion, and in particular their rotation vectors. A clustering algorithm is used to determine clusters of rotation vectors corresponding to main-chain segments that form possible dynamic domains. Domains are accepted for further analysis on the basis of the ratio of interdomain to intradomain fluctuation. Finally the residues involved in interdomain motion are identified. The methodology has been tested on citrate synthase and the M6I mutant of T4 lysozyme. In both cases new aspects to their conformational change are revealed, as are individual residues intimately involved in their dynamics. For citrate synthase, the β-sheet is found to be part of the hinging mechanism. In the case of T4 lysozyme, one of the four transitions in the pathway from the closed to the open conformation, furnished four dynamic domains rather than the expected two. This result indicates that the number of dynamic domains a protein possesses may not be a constant of the motion *(69)*.

Wriggers and Schulten *(70)* have presented an algorithm to identify and visualize the movements of rigid domains about common hinges in proteins. In comparing two structures, the method partitions a protein into domains of preserved geometry. The domains are extracted by an adaptive selection procedure using least-squares fitting. The user can maintain the spatial connectivity of the domains and filter significant structural differences (domain movements) from noise in the compared sets of atomic coordinates. The algorithm subsequently characterizes the relative movements of the found domains by effective rotation axes (hinges). The method is applied to several known instances of domain movements in protein structures, namely, in lactoferrin, hexokinase, actin, the extracellular domains of human tissue factor, and the receptor of human growth factor.

While a rigid-body matching can still be performed between the model structure and all other potential target structures in the family, such a procedure would need to be complemented by coalescing a large number of potential pairs of matches for the detection of each hypothetical hinge. Additionally, it is quite possible that each of the parts of the rigid body match contributing to such a hinge-bending motion may not contain a large enough number of matching C_α pairs from the model and the target proteins. Hence, the correct match may rank rather low in the list of potential matches. Certainly, carrying out a structural comparison *a priori* allowing hinge-bending, swiveling motions is a more desirable strategy. However, what is entailed below is performing extensive, fast, structural comparisons, allowing hinge-bending movements, where the motif is not predefined, and the locations of the hinges and the relative rotations between two adjoining domain (subdomain) parts are unknown. Carrying out such a task efficiently, avoiding the very time-consuming conformational space search, is a formidable task.

Thus, whereas there are numerous approaches to finding recurring substructural motifs in protein structures, where the structure recurs as a rigid body, there are few methods that can detect motifs allowing hinges between them. This scarcity in such automated methods is owing to the more difficult problem. We have recently developed and implemented an algorithm designed to accomplish such a goal *(71)*. Our algorithm is very fast. Comparison of two protein structures takes only seconds on a Silicon Graphics Indigo 2 R4400, 150 MHZ workstation. It carries out the comparisons in a manner independent of the sequence order of the amino acids on the chains, hence allowing a change in the directionality of the chains as well as insertions and deletions. Although currently implemented for one hinge, the approach is general, and several hinges can be straightforwardly incorporated. All internal rotational degrees of freedom are

allowed at the hinge. In particular, this is done in a highly efficient manner, avoiding the very time-consuming full conformational-space search noted previously. We are able to achieve such a performance because this robotics-based technique uses a hinge-based, internal coordinate reference frame. In a manner analogous to a robot moving its limbs or rotating and tilting its head, we allow molecular parts such as domains, subdomains, and loops to rotate around potential, preselected point-hinges. By allowing the parts' complete 3D rotations around a point, rather than the simpler case of rotation around a covalent bond, we implicitly take into account several consecutive, or nearby rotatable bonds. As domain motions are allowed (unlike in a rigid body superposition) two transformations are reported by the program. The first corresponds to the superposition of the first domain from the first protein onto the first domain from the second; the second transformation corresponds to the superposition of the second domain from the first protein onto the second domain from the second protein. The high efficiency of this algorithm allows the scanning of many potential hinge sites, ensuring finding all possible matchings.

The ability of the technique to detect similarities between molecules when these are considered articulated objects illustrates our inherent ability to handle automatically domain (subdomain; fragment) movements. Each of the domains moves as a rigid body, still maintaining its packing, with the deformation confined largely to the linking hinge regions. Our method generalizes and extends approaches to molecular similarity by considering molecules as articulated objects with predefined hinges. To illustrate the logic of the algorithm, consider the molecule as an ordered set of its C_α atoms. We may select a specific C_α atom and divide the molecule into two parts. The molecule is now considered as an object comprised of two rigid parts, connected by the selected C_α atom. This C_α is the hinge between the two parts. The two parts can swivel with respect to each other about the C_α atom. There is no limitation in the selection of the hinge site. It may be positioned anywhere in 3D space, and need not be specifically one of the C_α atoms of the molecule. It may be inside a loop, enabling two secondary structure elements to swivel with respect to each other, or it may be anywhere within the molecule. Alternatively, it may be positioned outside the molecule; for example, at the intersection of vectors describing secondary structure elements, extending the repertoire of allowed motions.

Thus, given a target molecule, represented by its C_α atom coordinates, and a database of molecules, each of which may possess a small number of hinges, we can find those molecules that, under an appropriate translation and rotation of the whole molecule, in addition to appropriate rotations at the hinges, will have a set of its C_α atoms superimposed on those of the target molecule. If the locations of the hinges in the model proteins are known, or predicted, they are incorporated in advance. If they are unknown, and no *a priori* assumptions about their locations can be made, the algorithm is fast enough to allow scanning of the molecules, with multiple potential locations. We give a short sketch of the major steps of the method for a model molecule having a single hinge:

1. Each of the rigid model molecules is represented by its transformation independent features (invariants). Cartesian reference frames are defined both at the invariant features and at the hypothetical hinge position. For each invariant, the rotation and translation between its reference frame and the hinge-based frame is stored in a hash table.
2. The target molecule is represented by its transformation-independent features as well, which are compared to the information stored in the hash table. Based on this comparison, potential hinge-based reference frame locations are detected and clustered. The high-scoring clusters of hinge-based reference frames generate potential transformations between the appropriate parts. The use of such a hash table for storing the transformation-independent features makes the comparison extremely efficient.

3. The potential matches are subsequently verified, by transforming the model onto the target, and evaluating the match.

Thus, it is easy to see why the algorithm is fast, and how it allows implicit scanning of all degrees of freedom between the rotating domains *without* searching the entire conformational space. It is only *after* the matches are found that the transformations are computed and applied to the entire domains. The output from this first stage are the hinge-based, substructural motifs superimposed on other locally similar proteins, with the domains rotated with respect to each other, optimally matching the detected protein homologue. Only after detection of such hinge-bending domain motions between otherwise similar substructural motifs, and the superpositioning of the corresponding domains, may be derived the residue interchanges at structurally analogous positions.

To give a flavor of the performance of the algorithm, it has been tested on approx 30 cases. The preprocessing time took between 0.10–0.83 s, and the recognition times varied from 16.86–148.58 s. The RMSDs ranged between 0.77–1.69 Å.

The approach described here utilizes structural similarity between locally or globally related structural pairs to detect hinge-bending motions. An alternate complementary approach has recently been shown to predict modes of motions *(72)*. The slowest modes are those resulting from the largest scale domain (subdomain) motions. This analysis does not require structural similarity, and hence in the absence of related structures, can be a method of choice for detection of potential, hypothetical, hinge-bending movements.

CONCLUSIONS

We have argued that the conformational alterations observed in the comparisons of open and closed structures are not the result of an induced fit binding mechanism. Rather, we have proposed that around the bottom of the folding funnels, flexible molecules exist in a range of conformational isomers. Hence, the conformation that is observed in the unbound molecule is the one most favorable for binding in the crystal under these conditions. Upon its binding to the growing crystal, the equilibrium shifts in its direction. On the other hand, the conformation observed in the bound state is the one most favorable for binding the ligand. Hence, upon binding the ligand, the equilibrium shifts in the direction of the conformer that is most complementary to the ligand. We have further argued that the two (or more) conformers that have been observed in the crystals represent only very few of the isomers that populate the bottom of the funnel. Geometrically similar isomers are near each other on the energy landscape. The examples we have listed illustrate conformers with a hinge-bending type of motion. Here, the molecular parts move with respect to each other as relatively rigid bodies. These structural changes represent low-energy transitions, and hence may take place in solution in the absence of the ligand.

We have divided our examples into two classes to highlight two major concepts. The first involves distinction according to the type of the structural part involved, namely, motions of fragments, of domains, and of subunits. The second is the motions in allosterically and non-allosterically regulated proteins.

Folding and Binding: Similar Hierarchical Processes

The illustration that different structural parts—whether short consecutive fragments of the polypeptide chain, or larger, independent structural entities such as domains, or a hydrophobic folding unit—manifest the same types of motions is consistent with the hierarchical notion of protein folding *(14)*. Furthermore, these, in turn, show similar types of conformational transition as those observed between subunits, which are not covalently connected. Taken together, it is again consistent with binding and folding being physically similar processes, conforming to the same physical laws. The only difference is that

in the fragment hinge-bending motions, or similarly in the case of the domains, the movements are constrained, due to the backbone connectivity, with allowed, and disallowed rotations. This is not the case for subunits. In the latter, subunit case a larger range of variability can be expected to exist. In this regard it is interesting to examine the case of the triosephosphate isomerase. The mutational analysis which has been carried out *(25)* has illustrated some clear preferences in selectivity for residues at, and near the hinges. While too-rigid, or bulky residue-combination, was unfavorable, overly flexible hinges were not favored as well.

Allostery

Most of the conformational changes in allosteric regulation involve subunit motions. Binding of the inhibitor favors the T state. Binding of the substrate or the activator favors the R state. Allostery can similarly be related to the concept of the funnels. The same principles and concept apply. However, as the system is larger, and hence more conformations are potentially possible, the funnels will be more complex. Furthermore, in contemplating allostery, we should consider binding funnels. In general, whereas the bottoms of the folding funnels are populated by ensembles of conformations of single protein molecules, the bottoms of the binding funnels are populated by an ensemble of complexed, bound conformations. For allosterically regulated proteins, there are, however, additional considerations. First, frequently allostery is observed in the quaternary molecular assembly, involving several subunits, as we have previously described. Second, the binding of the inducer may need to be considered, in addition to the substrate. Hence, allostery may involve large multi-molecular complexes, with higher-dimensional, very complicated funnels.

In summary, we have shown that the new view of protein folding can be extended to explain the fundamental concepts behind theories of induced fit and lock-and-key binding. Exploiting the similar nature of protein folding and binding processes, we show that the concept of the energy funnel is equally useful in understanding protein function.

Throughout this article, we have paid particular attention to the role played by conformational flexibility and structural fluctuations in enzyme activity. We have shown that a large variety of internal motions, extending over different time scales and with different amplitudes are involved in the catalytic cycle. Within these, we have focused particularly, though not entirely, on conformational changes upon substrate binding, which occur in enzymes consisting of two domains. These have a considerable effect on the catalytic activity of enzymes. Because conformational transitions involving hinge-bending motions are very frequent, and consequently functionally important, we have chosen to center on these in this work.

ACKNOWLEDGMENTS

We thank Dr. Jacob Maizel for encouragement and for helpful discussions. The research of H. Wolfson and R. Nussinov in Israel has been supported in part by grant number 95-00208 from BSF, Israel; a grant from the Israel Science Foundation administered by the Israel Academy of Sciences; the Magnet grant; the Ministry of Science grant; and 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 research of H. Wolfson is partially supported by the Hermann Minkowski-Minerva Center for Geometry at Tel Aviv University. 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 US Government.

REFERENCES

1. Fischer, E. (1894) *Ber. Dt. Chem. Ges.* **27,** 2985–2991.
2. Koshland, D. E., Jr. (1958) Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. USA* **44,** 98–123.
3. Bryngelson, J. D. and Wolynes, P. G. (1989) Intermediates and barrier crossing in a random energy model (with applications to protein folding). *J. Phys. Chem* **93,** 6902–6915.
4. Karplus, M. and Shakhnovitch, E. (1992) Protein folding: theoretical studies of thermodynamics and dynamics, in *Protein Folding* (Creighton T. ed.), W. H. Freeman & Sons, New York, pp. 127–195.
5. Karplus, M., Sali, A., and Shakhnovitch, E. (1995) Comment: kinetics of protein folding. *Nature* **373,** 664–665.
6. Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995) Navigating the folding routes. *Science* **267,** 1619–1620.
7. Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z., and Socci, N. D. (1995) Towards an outline of the topography of a realistic protein folding funnel. *Proc. Natl. Acad. Sci. USA* **92,** 3626–3630.
8. Baldwin, R. L. (1994) Matching speed and stability. *Nature* **369,** 183–184.
9. Baldwin, R. L. (1995) The nature of protein folding pathways: The classical versus the new view. *J. Biomolec. NMR* **5,** 103–109.
10. Dill, K. A. and Chan, H. S. (1997) From Levinthal to pathways to funnels. *Nature Struct. Biol* **4,** 10–19.
11. Karplus, M. (1997) The Levinthal paradox: yesterday and today. *Folding Design* **2,** S69–S75.
12. Lazaridis, T. and Karplus, M. (1997) "New view" of protein folding reconciled with the old through multiple unfolding simulations. *Science* **278,** 1928–1931.
13. Gruebele, M. and Wolynes, P. (1998) Satisfying turns in folding transitions. *Nature Struct. Biol.* **5,** 662–665.
14. Tsai, C.-J., Xu, D., and Nussinov, R. (1998) Protein Folding via Binding, and vice versa. *Folding & Design* **3,** R71–R80.
15. Gerstein, M., Lesk, A., and Chothia, C. (1994) Structural mechanism for domain movements in proteins. *Biochemistry* **33,** 6739–6749.
16. Gerstein, M. and Krebs, W. (1998) A database of macromolecular motions. *Nucleic Acids Res.* **26,** 4280–4290.
17. Derewenda, U., Brzozowski, A. M., Lawson, D. M., and Derewenda, Z. S. (1992) Catalysis at the interface: the anatomy of a conformational change in a triglyceride lipase. *Biochemistry* **31,** 1532–1541.
18. Brady, L., Brzozowski, A. M., Derewenda, Z. S., Dodson, E., Dodson, G., Tolley, S., et al. (1990) A serine protease triad forms the catalytic centre of a triaclglycerol lipase. *Nature* **343,** 767–770.
19. Bernstein, F., Koetzle, T., Williams, G., Meyer, E. J., Brice, M., Rodgers, J., et al. (1977) The Protein Data Bank: a computer based archival file for macromolecular structures. *J. Mol. Biol.* **112,** 535–542.
20. Peters, G. H., Toxvaerd, S., Olsen, O. H., and Svendsen, A. (1997) Computational studies of the activation of lipases and the effect of a hydrophobic environment. *Protein Eng.* **10,** 137–147.
21. Berg, O. G., Cajal, Y., Butterfoss, G. L., Grey, R. L., Alsina, M. A., Yu, B. Z., and Jain, M. K. (1998) Interfacial activation of triglyceride lipase from Thermomyces (Humicola) lanuginosa: kinetic parameters and a basis for control of the lid. *Biochemistry* **37,** 6615–6627.
22. Lowe, M. E. (1997) Colipase stabilized the lid domain of pancreatic triglycerate lipase. *J Biol. Chem.* **272,** 9–12.
23. Iwata, S., Kamata, K., Yoshida, S., Minowa, T., and Ohta, T. (1994) T and R states in the crystals of bacterial L-lactate dehydrogenase reveal the mechanism for allosteric control. *Struct. Biol.* **1,** 176–185
24. Klimasaukas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) HhI methyltransferase flips its target base out of the DNA helix. *Cell* **76,** 357–369.
25. Sun, J. and Sampson, N. S. (1998) Determination of the amino acid requirements for a protein hinge in triosephosphate isomerase. *Protein Sci.* **7,** 1495–1505.
26. Mainfroid, V., Terpstra, P., Beauregard, M., Frere, J. M., Mande, S. C., Hol, W. G. J., Martial, J. A., and Goraj, K. (1996) 3 HTIM mutants that provide new insight on why TIM is a dimer. *J. Mol. Biol* **257,** 441–456.
27. Rietveld, A. W. M. and Ferreira, S. T. (1998) Kinetics and energetics of subunit dissociation/unfolding of TIM: the importance of oligomerization for conformational persistence and chemical stability of proteins. *Biochemistry* **37,** 933–937.
28. Schliebs, W., Thanki, N., Eritja, R., and Wierenga, R. (1996) Activate-site properties of monomeric triosephosphate isomerse

(MONOTIM) as deduced from mutational and structural studies. *Protein Sci.* **5,** 229–239.
29. Talent, J. M., Zvaigzne, A. I., Agrawal, N., and Gracy, R. W. (1997) Effect of active-site modification on the terminal marking deamidation of triosephosphate isomerase. *Arc. Biochem. Biophys* **340,** 27–35.
30. Derreumaux, P. and Schlick, T. (1998) The loop opening/closing motion of the enzyme triosephosphate isomerase. *Biophys. J.* **74,** 72–81.
31. Auerbach, G., Huber, R., Grattinger, M., Zaiss, K., Schurig, H., Jaenicke, R., and Jacob, U. (1997) Closed structure of phosphoglycerate kinase from Thermotoga maritima reveals the catalytic mechanism and determinants of thermal stability. *Structure* **5,** 1475–1483.
32. Bernstein, B. E., Michels, P. A. M., and Hol, W. G. J. (1997) Synergistic effects of substrate-induced conformational changes in phosphoglycerate kinase activation. *Nature* **385,** 275–278.
33. Cheung, C. W. and Mas, M. T. (1996) Substrate-induced conformational-changes in yeast 3-phosphoglycerate kinase monitored by fluorescence of single tryptophan probes. *Protein Sci.* **5,** 1144–1149.
34. Ritcovonsovici, M., Mouratou, B., Minard, P., Desmadril, M., Yon, J. M., Andrieux, M., Leroy. E., and Guittet, E. (1995) Role of the C-terminal helix in the folding and stability of yeast phosphoglycerate kinase. *Biochemistry* **34,** 833–841.
35. Adams, B., Fowler, R., Hudson, M., and Pain, R. H. (1996) The role of the C-terminal lysine in the hinge bending mechanism of yeast phosphoglycerate kinase. *Febs Lett.* **385,** 101–104.
36. Pecorari, F., Minard, P., Desmadril, M., and Yon, J. M. (1993) Structure and functional complementation of engineered fragments from yeast phosphoglycerate kinase. *Protein Eng.* **6,** 313–325.
37. Pecorari, F., Guilbert, C., Minard, P., Desmadril, M., and Yon, J. M. (1996) Folding and functional complementation of engineered fragments from yeast phosphoglycerate kinase. *Biochemistry* **35,** 3465–3476
38. Parker, M. J., Spencer, J., Jackson, G. S., Burston, S. G., Hosszu, L. L.P., Craven, C. J., Waltho, J. P., and Clarke, A. R. (1996) Domain behavior during the folding of a thermostable phosphoglycerate kinase. *Biochemistry* **35,** 15740–15752.
39. Sherman, M. A., Beechem, J. M., and Mas, M. T. (1995). Probing intradomain and interdomain conformational-changes during equilibrium unfolding of phosphoglycerate inase–fluorescence and circuilar-dichroism study of tryptophan mutants. *Biochemistry* **34,** 13934–13942.
40. Lillo, M. P., Szpikowska, B. K., Mas, M. T., Sutin, J. D., Beechem, J. M. (1997) Real-time measurement of multiple intramolecular distances during protein folding reactions: a multisite stopped-flow fluorescence energy-transfer study of yeast phosphoglycerate kinase. *Biochemistry* **36,** 11273–11281.
41. McPhillips, T. M., Hsu, B. T., Sherman, M. A., Mas, M. T., and Rees, D. C. (1996) Structure of the R65Q mutant of yeast 3-phosphoglycerate kinase complexed with Mg-AMP-PNP and 3-phospho-D-glyceratefe. *Biochemistry* **35,** 4118–4127
42. Schulz, G. E., Muller, C. W., and Diederichs, K. (1990) Induced-fit movements in adenylate kinase. *J. Mol. Biol.* **213,** 627–630.
43. Schlauderer, G. J., Proba, K., and Schulz, G. E. (1996) Structure of A Mutant Adenylate kinase ligated with an ATP-analog showing domain closure over ATP. *J. Mol. Biol.* **256,** 223–227.
44. Muller, C. W, Schlauderer, G. J., Reinstein, J., and Schulz, G. E. (1996) Adenylate kinase motion during catalysis: An energetic counterweight balancing substrate-binding. *Structure* **4,** 147–156.
45. Zhang, H. J., Sheng, X. R., Pan, X. M., and Zhou, J. M. (1997) Activation of adenylate kinase by denaturants is due to the increasing conformational flexibility at its active sites. *Bioch. Biophys. Res. Comm.* **238,** 382–386.
46. Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., and Bax, A. (1992) Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science* **256,** 632–638.
47. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping: A mechanism for oligomer assembly. *Protein Sci.* **4,** 2455–2468.
48. Baker, P. J., Britton, K. L., Engel, P. C., Farrants, G. W., Lilley, K. S., Rice, D. W., and Stillman, T. J. (1992) Subunit assembly and active site location in the structure of glutamate dehydrogenase. *Proteins* **12,** 75–86.
49. Kumar, S., Tsai, C. J., Ma, B., and Nussinov, R. (1999) Contribution of salt bridges toward

protein thermostability. *J. Biomolecular Struct. & Dynamics.* In press.
50. Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., and Sigler, P. B. (1994) The crystal structure of the bacterial chaperonin GroEL at 2.8Å. *Nature* **371,** 578–586.
51. Stevens, R. C., Gouaux, J. E., and Lipscomb, W. N. (1990) Structural consequences of effector binding to the T state of aspartate carbamoyltransferase: crystal structures of the unligated and ATP- and CTP-complexed enzymes at 2.6Å resolution. *Biochemistry* **29,** 7691–7701.
52. Sakash, J. B. and Kantrowitz, E. R. (1998) The N-terminus of the regulatory chain of *Escherichia coli* aspartate transcarbamoylase is important for both nucleotide binding and heterotropic effects. *Biochemistry* **371,** 281–288.
53. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. (1996) Crystal structure of the lactose operon repressor and its complexes with the DNA and inducer. *Science* **271,** 1247–1254.
54. Milburn, M. V., Prive, G. G., Milligan, D. L., Scott, W. G., Yeh, J., Jancarik, J., Koshland, D. E., and Kim, S.-H. (1991) Three dimensional structures of the ligand-binding domain of the bacterial aspartate receptor with and without a ligand. *Science* **254,** 1342–1347.
55. Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1995) Structure of *Bam*HI endonuclease bound to DNA: partial folding and unfolding on DNA binding. *Science* **269,** 656–663.
56. Standfield, R. L., Takimoto-Kanimura, M., Rini, J. M., Profy, A. T., and Wilson, I. A. (1993) Major antigen-induced domain rearrangements in an antibody. *Structure* **1,** 83–93.
57. Herron, J. N., He, X. M., Ballard, D. W., Blier, P. R., Pace, P. E., Bothwell, A. L., Voss, E. Jr., and Edmundson, A. B. (1991) An autoantibody to single-stranded DNA: comparison of the three-dimensional structures of the unliganded Fab and a deoxynucleotide-Fab complex. *Proteins* **11,** 159–175.
58. Rini, J. M., Stanfield, R. L., Stura, E. A., Salinas, P. A., Profy, A. T., and Wilson, I. A. (1993) Crystal structure of a human immunodeficiency virus type 1 neutralizing antibody 50.1, in complex with its V3 loop peptide antigen. *Proc. Natl. Acad. Sci. USA* **90,** 6325–6329.
59. Marmorstein, R. and Harrison, S. C. (1994) Crystal structure of a PPR1-DNA complex: DNA recognition by proteins containing a Zn^2Cys^6 binuclear cluster. *Gen. Dev.* **8,** 2504–2412.
60. Marques, O. and Sanejouand, Y. H. (1995) Hinge-bending motion in citrate synthase from normal-mode calculations. *Proteins* **23,** 557–560.
61. Ikura, T. and Go, N. (1993) Normal-mode analysis of mouse epidermal growth-factor: characterization of the harmonic motion. *Proteins* **16,** 423–436.
62. Guilbert, C., Pecorari, F., Perahia, D., and Mouawad, L. (1996) Low-frequency motions in phosphoglycerate kinase: a normal mode analysis. *J. Chem. Phys.* **204,** 327–336.
63. Hayward, S., Kitao, A., and Berendsen, H. J. C. (1997) Model-free methods of analyzing domain motions in proteins from simulation: a comparison of normal mode analysis and molecular dynamics simulation of lysozyme *Proteins* **27,** 425–437.
64. deGroot, B. L., Hayward, S., vanAalten, D. M. F., Amadei, A., and Berendsen, H. J. C. (1998) Domain motions in bacteriophage T4 lysozyme: a comparison between molecular dynamics and crystallographic data. *Proteins* **31,** 116–127.
65. vanAalten, D. M. F., Grotewold, E., and Joshua Tor, L. (1998) Essential dynamics from NMR clusters: dynamic properties of the Myb DNA-binding domain and a hinge-bending enhancing variant. *Methods-A Companion To Methods In Enzymology* **14,** 318–328.
66. Alexov, E. and Atanasov, B. (1995) Selective absorption of radio-frequency energy due to collective motion of charged domain: case of lysozyme crystal. *J. Biomol. Struct. Dyn.* **13,** 219–228.
67. Miura, N., Hayashi,Y., and Mashimo, S. (1996) Hinge-bending deformation of enzyme observed by microwave dielectric measurement. *Biopolymer* **39,** 183–187.
68. Chandra, N. R., Muirhead, H., Holbrook, J. J., Bernstein, B. E., Hl, W. G. J., and Sessions, R. B. (1998) A general method of domain closure is applied to phosphoglycerate kinase and the result compared with the crystal structure of a closed conformation of the enzyme. *Proteins* **30,** 372–380.
69. Hayward, S. and Berendsen, H. J. C. (1998) Systematic analysis of domain motions in

proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* **30,** 144–154.

70. Wriggers, W. and Schulten, K. (1997) Protein domain movements: detection of rigid domains and visualization of hinges in comparisons of atomic coordinates. *Proteins* **29,** 1–14.
71. Verbitsky, G., Nussinov, R., and Wolfson, H. (1999) Structural comparison allowing hinge bending, swiveling motions. *Proteins* **34,** 232–254.
72. Bahar, I., Erman, B. Haliloglu, T., and Jernigan, R. L. (1997) Identification of cooperative motions and correlated structural elements in coarse-grained proteins: application to T4 lysozyme. *Biochemistry* **36,** 13512–13523.