Published in final edited form as:

Curr Protein Pept Sci. 2009 April ; 10(2): 116–127.

Frameworks for Understanding Long-Range Intra-Protein Communication

M. J. Whitley¹ and A. L. Lee^{1,2}

¹Department of Biochemistry & Biophysics, School of Medicine, University of North Carolina, Chapel Hill, NC 27599

²Division of Medicinal Chemistry & Natural Products, School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599

Abstract

The phenomenon of intra-protein communication is fundamental to such processes as allostery and signaling, yet comparatively little is understood about its physical origins despite notable progress in recent years. This review introduces contemporary but distinct frameworks for understanding intra-protein communication by presenting both the ideas behind them and a discussion of their successes and shortcomings. The first framework holds that intra-protein communication is accomplished by the sequential mechanical linkage of residues spanning a gap between distal sites. According to the second framework, proteins are best viewed as ensembles of distinct structural microstates, the dynamical and thermodynamic properties of which contribute to the experimentally observable macroscale properties. Nuclear magnetic resonance (NMR) spectroscopy is a powerful method for studying intra-protein communication, and the insights into both frameworks it provides are presented through a discussion of numerous examples from the literature. Distinct from mechanical and thermodynamic considerations of intra-protein communication are recently applied graph and network theoretic analyses. These computational methods reduce complex three dimensional protein architectures to simple maps comprised of nodes (residues) connected by edges (inter-residue “interactions”). Analysis of these graphs yields a characterization of the protein’s topology and network characteristics. These methods have shown proteins to be “small world” networks with moderately high local residue connectivities existing concurrently with a small but significant number of long range connectivities. However, experimental studies of the tantalizing idea that these putative long range interaction pathways facilitate one or several macroscopic protein characteristics are unfortunately lacking at present. This review concludes by comparing and contrasting the presented frameworks and methodologies for studying intra-protein communication and suggests a manner in which they can be brought to bear simultaneously to further enhance our understanding of this important fundamental phenomenon.

Keywords

proteins; intra-protein communication; allostery; conformational equilibrium; NMR spectroscopy; side-chain dynamics; communication pathways; network topology

Introduction

If asked to give a brief history of efforts to understand the foundations and consequences of long-range intra-protein communication, most biological scientists would begin with

investigations of the oxygen binding properties of hemoglobin. Indeed, the phenomenon of cooperative oxygen binding to hemoglobin launched a major scientific effort to understand how events at one location in a protein can influence events at other sites located many angstroms away in the same protein. This effort has spanned many decades and continues even today. In the latter half of the twentieth century, several now famous models were developed to explain the cooperative nature of ligand binding by proteins. The most famous two are known as the KNF and MWC models, and these were very successful at explaining basic features of cooperative transitions for various protein systems. Since their development, however, numerous experimental and theoretical observations have necessitated the development of increasingly sophisticated frameworks for explaining both how (mechanism) and why (energetics) intra-protein communication exists. This review article has several purposes. The first is to briefly explain the salient features of the two classical models of intra-protein communication, and then to introduce modern variations by drawing on the relevant recent literature. Secondly, the increasingly recognized role of protein dynamics in modulating protein function will be discussed, with specific attention given to the insights gained via high resolution nuclear magnetic resonance (NMR) spectroscopy and how these insights can be understood in terms of the two classic models for allosteric communication. The final purpose of this paper is to highlight the analysis of intra-protein interactions using concepts from information and network theory. The novelty of information and network theoretic analyses applied to biological problems reveals yet again that scientific questions may be answered most completely when examined from several different points of view. This review will serve a useful purpose if information and network theorists and experimental biological scientists can be introduced to the language and methods of the two respective fields and thereby facilitate collaborations to increase understanding of the causes and consequences of intra-protein communication.

I. Mechanisms of Intra-Protein Communication: Communication Models Then and Now

The KNF or Sequential Model of Signal Propagation

An article by Koshland published in 1958 ushered in a new framework for thinking about ligand binding by proteins. In that seminal article, Koshland suggested that ligand binding actually occurs by an “induced fit” mechanism [1]. This idea holds that the 3D architectures of ligand binding sites on proteins are not perfectly complementary to the intended ligand (i.e., that binding does not occur through a lock-and-key fit) and that ligand binding causes a spatial reconfiguration of the binding site so as to maximize favorable interactions with the ligand. The concept of induced fit was later incorporated into the KNF or sequential model of cooperativity in multimeric proteins [2], which posits that conformational changes caused by induced-fit ligand binding to one subunit propagate to the next subunit, making it now slightly more thermodynamically favorable for ligand to bind to that subunit. This process repeats itself sequentially until ligand is bound to all subunits. What this model does not state, however, is the physical mechanism by which ligand binding in the first subunit results in slightly increased binding affinity for the second unit. One possibility for transmitting the presence of a bound ligand to remote sites is that the conformational rearrangements induced by the ligand at its binding site can themselves induce conformational rearrangements of the next shell of the protein, which then induce further rearrangements at even greater distances from the binding site and so on until the free energy of binding has been exhausted. The key feature of this type of long-range communication is that the transmission of the signal (in this case, that a ligand has bound somewhere on the protein) happens sequentially in both space and time. This mode of intra-protein communication is mechanistic in the most literal sense of the word – the physical movement of certain amino acid residues causes them to impact other residues, which themselves strike additional residues, and so the signal is propagated through a series of

physical interactions occurring sequentially in both space and time. This idea is displayed schematically in Fig. (1a).

Application of the Sequential/Mechanical Model to Membrane Proteins

Several interesting examples of this mode of internal communication are to be found in the recent literature. A single-molecule kinetic study of the acetylcholine receptor revealed the temporal nature of the allosteric conformational transition linking agonist binding in the extracellular domain to increased ion permeability in the transmembrane channel [3]. The researchers measured the channel opening and closing rates for a series of mutants at numerous positions in the transmembrane channel and used this information to calculate a Φ value for each residue, which is the slope of a log-log plot of the opening rate versus the equilibrium constant of the open/closed equilibrium. The Φ value can be interpreted as the relative timing of the movement of a particular residue during the allosteric transition. Based on their Φ values and previously published Φ values for other regions of the receptor, the authors determined that agonist binding in the extracellular space is linked to increased ion permeability through the transmembrane channel by a series of clustered residue motions, with the first such motions happening nearest to the agonist binding site and other clusters generally moving in order of increasing distance from the agonist site. In this way, “a Brownian conformational cascade” of residue motions links agonist binding to changed ion permeability in the transmembrane pore [3]. Mechanical coupling is also suspected to be the method of communication between the activation gate and the selectivity filter of the voltage-activated potassium channel (Kv). Previous studies of these channels have delineated gating-sensitive residues, those residues for which mutation alters the channel open/closed equilibrium, and it was noticed that these residues cluster not only near the activation gate and selectivity filter elements of Kv, but also on a pathway connecting these elements [4]; double mutant cycle analyses of these positions revealed that members of this pathway are energetically coupled [5,6]. Against this background, Sadvovsky and Yifrach recently examined whether higher order coupling relationships exist among combinations of on-path residues [7]. They discovered that coupling energies along the proposed allosteric trajectory increase with the order of the coupling and that couplings among path residues are reduced or abolished when higher order interaction partners elsewhere on the path are removed. These results suggest that all residues along the proposed communication pathway are strongly coupled and that all are essential for efficient signal transmission between activation gate and sensitivity filter. Because the free energies of all coupling relationships measured were positive (implying that the underlying interactions are stronger in the channel’s closed state), the researchers suggest that ion flow is physically prevented by a series of interactions along the pathway in the closed state and that activation leads to the propagation of slight conformational changes down the allosteric trajectory, resulting in weakened interactions among on-path residues and allowing ions to flow through the channel. Results from studies of these two membrane channel systems clearly indicate that deformations caused by perturbations are capable of propagating over long distances in proteins and that these structural rearrangements can influence functional properties of the systems in question.

The MWC Model and Conformational Equilibria

Another framework for understanding long-range intra-protein communication looks beyond the mechanical deformation of a system in response to perturbation (a macroscopic phenomenon) and instead focuses on a protein’s microscopic thermodynamic characteristics. The initial foundation of this idea is the MWC or concerted model of cooperativity, again developed to explain oxygen binding by hemoglobin [8]. One of the key foundations of this model is that both subunit conformations (classically, the T and R states) exist and are in thermal equilibrium with each other even in the total absence of ligand. Both conformations are capable of binding ligand, albeit with noticeably different affinities. Rather than inducing a change in

the other subunits toward the more favorable ligand binding conformation upon ligand binding to the first subunit, the MWC model holds that ligand binding stabilizes a particular subunit in the high affinity form, and this stabilization simply shifts the equilibrium of the whole oligomeric protein decidedly toward the high affinity state. In short, ligand binding does not *cause* a conformational change, it simply alters a pre-existing equilibrium to distinctly favor one conformation over the other. This idea is displayed graphically in Fig. (1b).

Modern versions of this pre-existing equilibrium theory come in many flavors, but they all have several basics in common: (i) a protein conformation is not static; rather, it constantly populates a range of distinct conformations according to the Boltzmann distribution, (ii) ligands bind to a small subset of the conformations present in the ensemble, stabilizing them such that the probability of remaining in those conformations increases (assuming positive cooperativity), and (iii) the entire ensemble eventually becomes enriched in the ligand-bound, stabilized conformations. Macroscopically, the presence of ligand appears to “cause” structural rearrangements resulting in a new average ligand-bound conformation, but in fact the ligand has only made the subset of binding-competent conformations, which were always present to some degree, much more likely than the host of other possible conformations. If the selected conformations have a suitable spatial geometry at a binding site elsewhere in the protein, then those conformations will have different functional characteristics than the numerous other possible conformations originally present in the ensemble, and the original ligand will by definition be an allosteric regulator.

Investigations of the MWC Model and Pre-Existing Structural Equilibria

The central tenet of the MWC model, namely that a protein exists in an equilibrium between high- and low-affinity states in the absence of ligand, has recently received clear experimental support in the multimeric allosteric enzyme aspartate transcarbamoylase (ATCase). The D236A mutant of that enzyme lacks a key inter-subunit salt bridge that stabilizes the T-state (low ligand affinity) of the enzyme. Solution small angle X-ray scattering curves were recorded as a function of temperature for the WT and D236A forms of ATCase [9]. The WT curves show no variation over a wide range of temperatures (4–60°C); if an equilibrium between two states exists for the WT, the energy barrier for interconversion must be too high for interconversion to occur in the temperature range studied. For D236A ATCase, however, noticeable changes are obvious even over the narrower temperature range of 6–45°C. These changes are reversible with increasing or decreasing temperature and provide experimental evidence that this mutant exists in different conformations in dynamic equilibrium with each other and that the energy barrier between them is modest enough to allow noticeable population shifts within the temperature range studied. A subsequent NMR study of the binding of substrates, substrate analogues, and allosteric effectors to WT ATCase was again able to detect the presence of two distinct states in dynamic equilibrium with one another and extended the analysis of this system by using NMR ligand titrations to directly calculate the equilibrium constant for interconverting between the two states [10]. The NMR and SAXS studies of ATCase provide clear support to the MWC/pre-existing equilibrium model of intra-protein communication, and since its introduction, numerous studies have expanded the basic MWC model to consider the energetics of large numbers of possible protein conformations [11].

Computational techniques lend themselves to the study of ensemble-based phenomena, since they can analyze the properties of large numbers of protein structures in a tractable period of time. Among numerous other computational protein ensemble studies, a notable computational investigation that examines the three properties of the pre-existing equilibrium/ensemble method of long distance communication is the study of the effects of binding between the antibody D1.3 and hen egg white lysozyme (HEWL) [12]. In this study, a huge ensemble of theoretical HEWL structures was generated by varying different locations and extents of

local unfolding in the protein. The energies and probabilities of these states relative to the globally unfolded protein were calculated, and the effects of antibody binding on the distribution of states in the ensemble were then explored. In the absence of the antibody, the HEWL conformational ensemble contains many substates in which HEWL helix F is locally unfolded (i.e., has low stability). Helix F is also the major constituent of the D1.3 binding interface. Upon binding to D1.3, a major redistribution of the HEWL structural ensemble was observed, and helix F is folded in most of the new conformations. This observation agrees with the ensemble view of cooperativity. Helix F is stabilized by antibody binding, which makes populating conformational substates with an unfolded helix F much less probable. Antibody binding has selected a small subset of the original ensemble and selectively enriched the resulting ensemble in those substates. Notice too that this is not an example of an allosteric transition in the classical sense (i.e., the biological interaction between the antibody and enzyme was not designed to redistribute the HEWL ensemble in order to change a functional property of the enzyme, unlike the above example of ATCase). This emphasizes that long-range intra-protein communication can occur and be observed outside of a purely functional context and begs the questions as to how nature harnesses the potential for long-range communication (fundamental phenomenon) to enable allostery (specifically functional consequences).

In this section, we have summarized the development of two main theories of long-range cooperativity in proteins (the sequential and ensemble models) and how these models have been adapted since their development to explain the numerous experimental and theoretical examples of communication between distal sites in proteins. The former suggests that signals can be transmitted throughout proteins by a series of small conformational changes induced upon perturbation. Indeed, it is well known that very slight conformational adjustments can greatly affect protein function [13]. In comparison, the latter framework suggests that it is the microscopic properties of the protein (local conformational fluctuations and stabilities) that lead to a small subset of conformations gaining dominance over the rest in response to perturbation. The ensemble model might be more generally applicable than the sequential model. Numerous computational studies of protein cooperativity from the ensemble perspective [14–16] have repeatedly shown that the conformational substates selected by a certain perturbation depend on the energetics of coupling between residues. That the strength of these couplings is not necessarily correlated with the spatial proximity of the coupled residues reveals that sites communicating with one another need not necessarily be linked by a pathway of physical interactions. Finally, readers are referred to recent work by Cui and coworkers, who used MD simulations to discriminate between mechanical propagation *vs.* population shifting in the bacterial response regulator CheY and discuss similar issues raised here [17,18]. In the next section, we further examine conformational dynamics across multiple time scales, introduce how NMR spectroscopy can characterize protein motions, and discuss how protein dynamics might fit into the prevailing theories of long-range intra-protein communication.

II. Examining the Important Role of Protein Dynamics in Intra-Protein Communication

Internal Dynamics as a Possible Allosteric Regulator

It is against the background of the conformational ensemble that the internal dynamics of proteins becomes important when considering their possible role in intra-protein communication. The classical KNF and MWC models assume that, whatever the mechanism of the transition, the final conformational state of a macromolecule after an allosteric transition will be structurally distinct from the molecule's original conformation. However, by considering the role of internal dynamics, or random fluctuations around the mean conformation, it is possible to develop a framework for intra-protein communication that does

not depend on observable structural rearrangements [19]. This notion is based on the already familiar idea that the macroscopically observable structure of a protein is actually a dynamic ensemble of interconverting structures characterized by conformational fluctuations of varying amplitudes occurring with probabilities determined by the Boltzmann distribution. As already discussed above, ligand binding to one particular conformation in the ensemble can stabilize that conformation, increasing its lifetime and eventually enriching the ensemble in that conformation. The origin of this population redistribution is enthalpic in nature and likely results in observable structural differences. However, the possibility also exists that, rather than redistributing the conformational ensemble around a new average structure, the conformational ensemble remains centered around the original structure but now simply has different characteristics. This concept is displayed graphically in Fig. (2). Internal dynamics influences the conformational entropy of the system. If a perturbation alters the system's internal dynamics, the characteristic width of its conformational distribution will change, making some conformational substates more or less probable while retaining the same average macroscopic conformation. If the conformational substates affected are functional, then the perturbation of internal dynamics will have caused an allosteric response in the absence of gross structural rearrangements [20].

Using NMR Spectroscopy to Characterize Flexibility & Dynamic Propagation in the ps-ns Time Regime

Solution-state nuclear magnetic resonance spectroscopy is the leading technique for characterizing protein conformational flexibility with atomic resolution. To characterize the flexibility of protein backbones in the ps-ns time regime, standard NMR relaxation experiments are employed to measure the T_1 and T_2 relaxation time constants (the inverses of the R_1 and R_2 relaxation rate constants) and the heteronuclear nuclear Overhauser enhancement (NOE) for proteins enriched with ^{15}N at the amide nitrogen position (for recent reviews specifically addressing the analysis of these measurements, see [21,22]). Briefly, 2D ^1H - ^{15}N correlation spectra are recorded for a series of increasingly longer relaxation delays, and the intensities of the resonances (one for each amide group in the protein) as a function of relaxation delay are fit to a single exponential decay, yielding the relaxation time (or rate) constants. The T_1 , T_2 , and NOE data for each residue can then be analyzed using the Lipari-Szabo "model free" formalism [23,24] to yield the generalized order parameter S^2 and the correlation time τ_c . Rigid amide bond vectors have S^2 values of 1, while those able to reorient isotropically have S^2 values of 0. Because of its independence from physical models, the Lipari-Szabo formalism only gives information about the amplitude and time scale of motions experienced by a particular bond vector; it provides no information about the directionality or correlation of any particular motions with regard to other internal motions. A similar type of dynamics analysis can be performed to analyze the ps-ns dynamics of methyl-bearing side chains in proteins uniformly enriched with ^{13}C and 60% randomly enriched with ^2H [25–27]. This type of analysis yields order parameters and correlation times for the methyl symmetry axis in side-chain methyl groups.

Numerous recent studies have examined the ability of various perturbations to alter or modulate the fast conformational dynamics of protein backbones and side chains. Because modern high-resolution NMR spectroscopy yields site-specific information, responses to perturbations can be mapped to a protein's structure and thus reveal pathways of propagation. Lee and coworkers have examined the dynamical consequences of two different types of perturbation to the small, classically non-allosteric serine protease inhibitor eglin c. Through the application of ^{15}N and ^2H NMR relaxation experiments to eglin c under neutral and acidic solution conditions (pH 7 vs. pH 3), it was shown that under acidic conditions the protein's backbone rigidified slightly, while the rigidity of the methyl-bearing side chains in the ps-ns time regime increased by nearly 15% on average when compared to the dynamics parameters obtained at neutral pH

[28]. Also, the global stability of the protein is reduced from 6.1 kcal mol⁻¹ at pH 7 to only 2.7 kcal mol⁻¹ at pH 3. The authors note that protein destabilization is often attributed to lost favorable internal interactions and that this is assumed to result in a more flexible conformation, which is the opposite of what was observed for eglin c. Can these results be understood in terms of the conformational properties of the native state ensemble? Lowering the solution pH will increase proton binding by the protein and thus select for those microstates in the conformational ensemble that are competent for binding protons. If these proton binding conformations are distinct from the prevailing conformations at neutral pH, it is quite likely that they will have different free energies, exactly as was measured in this study. Furthermore, the decrease in stability at lower pH can also be rationalized in terms of ensemble characteristics. Increases in backbone and side-chain rigidity at pH 3 mean that the backbone and side-chains are sampling a narrower conformational distribution (Fig. (2b)), which results in a loss of conformational entropy. All other things being equal, a loss of conformational entropy will result in an overall destabilization of the protein.

Additional studies of eglin c have examined the effects of point mutations on internal dynamics. A study of two valine-to-alanine mutations (V14A, V54A) in eglin c's hydrophobic core revealed that these cavity-creating mutations result in spatially contiguous networks of altered side-chain dynamics extending up to 13 Å from the site of mutation, with the V54A network responses being rigidifications and the V14A side chains showing both increases and decreases in rigidity [29]. Equally significant is the lack of noticeable structural changes caused by the mutations, as assessed by measurements of residual dipolar couplings. It is worth repeating that eglin c is not a classically allosteric protein, but the fact that a perturbation (mutation) causes the protein's network of internal dynamics to change without a noticeable structural change is reminiscent of the long-range communication model of Fig. (2b). However, the fact that altered dynamics followed distinct pathways through the structure is suggestive of the possibility of an internal communication mechanism analogous to Fig. (1a). A follow-up study of additional eglin c core mutants revealed similar results. Some mutations resulted in distinct pathways of altered ps-ns side-chain dynamics, while other mutants resulted in more disperse, less cohesive responses [30]. One of these mutants, V34A, resulted in a slight structural change at a side chain located about 11 Å from the site of mutation. This side chain (V18) shows a shift in its major χ_1 rotamer in response to the mutation at V34. Most significant about this side-chain structural shift is that it takes place against the background of a network of altered side-chain dynamics and the absence of any gross backbone structural changes. Considering these contiguous pathways of altered dynamics to be a possible manifestation of the mechanical/sequential mode of long-range communication and thus viewing the rotameric shift of a distant side chain in response to changed internal dynamics as an allosteric-like response is to suggest that all globular proteins are fundamentally capable of long-distance communication and allosteric-like phenomena [31]. There are also other examples in the recent literature of mutation-dependent perturbations to internal dynamics. Four different mutations to calmodulin complexed with peptide result in four different patterns of altered internal dynamics. Mutant D95N hardly perturbs calmodulin's side-chain dynamics at all, whereas the pseudosymmetric mutation D58N results in significant long-range perturbation of methyl side-chain dynamics; additionally, the M124 mutation results only in local changes to side-chain dynamics, whereas E84K shows long-range propagation of altered dynamics [32] (for additional "perturbation-response" studies, see [33–36]). In a similar vein, a computational perturbation-response investigation of the open/closed transition in DNA and RNA polymerases has also revealed that a network of dynamically-important residues is necessary for the efficient coordination of the normal mode for opening and closing [37].

Other NMR relaxation studies have probed changes in internal dynamics in the ps-ns time regime in response to perturbations from ligand binding. The PDZ family of protein-protein interaction modules has been well characterized in terms of internal dynamics. A study of the

dynamics of the methyl-containing side chains of PDZ2 from human tyrosine phosphatase 1E in the presence and absence of a peptide target conducted via ^{15}N and ^2H NMR relaxation measurements [38] revealed that, when PDZ2 binds to a target peptide, many of the methyl-containing side chains involved in binding the peptide rigidify significantly, but also that there are changes in side-chain dynamics at positions not directly involved in peptide binding; furthermore, these distal residues undergoing altered dynamics radiate from the binding site to two distal surfaces (Fig. (3)). It has been suggested that these pathways connecting the peptide binding site to distal surfaces serve as a series of linked interactions that transduce the energy of peptide binding throughout the rest of the protein. This idea is supported by statistical and computational analyses that detected similar energetic pathways in related PDZ domains [39–41]. Another study of PDZ2 examined the ability of a member of the aforementioned energetic pathway to modulate PDZ2's internal dynamics upon mutation in comparison to off-pathway residues. The on-path mutation (I20F) resulted in extensive changes to the ps-ns side-chain dynamics of the protein in a manner reminiscent of the effects of peptide binding, whereas the off-path residues had minimal effects on dynamics elsewhere in the protein [42]. That mutation and peptide binding result in altered dynamics across similar pathways through the protein suggests that these pathways are physically relevant to long-range energy dispersion within the protein.

Relationships between conformational dynamics and function have also been noted for Pin1, a peptidyl-prolyl isomerase. Upon binding a phosphopeptide substrate, an internal pathway of hydrophobic side-chain rigidifications is observed that connects the isomerase active site with the interface between Pin1's two domains [43]. A later study revealed a link between protein dynamics and Pin1 binding affinity. Upon peptide binding, a flexible loop thought to be important for substrate recognition and binding was shown to rigidify across multiple time scales. The authors deleted a single residue in the loop (while still maintaining the chemical groups necessary for peptide interaction) and noticed that both the inherent flexibility of the loop and Pin1's ability to bind its substrate decreased [44]. This provides a clear link between Pin1's conformational entropy (the inherent flexibility of regions of the protein) and its function (ability to bind physiological interaction partners). Modulating the dynamics of the protein therefore has an allosteric affect on protein function. The idea that protein conformational dynamics can play a large role in protein/ligand interactions was recently demonstrated clearly and elegantly through a study of the changes in ps-ns dynamics of calmodulin upon binding a host of relevant peptide targets [45]. By using changes in side-chain ps-ns dynamics as a proxy for changes in conformational entropy, NMR and thermodynamic measurements revealed that the change in protein conformational entropy is linearly related to the total change of entropy of protein/peptide binding, and furthermore that the change in conformational entropy is a major and sometimes dominant contributor to the overall entropy and thus the overall free energy of binding. This suggests that modifying protein dynamics and thus conformational entropy can be used to tune a protein's preference for various binding partners.

Collectively, the studies discussed above reveal two features about the role of ps-ns dynamics as a method of internal communication in proteins. First, the studies reveal that many perturbations to proteins result in contiguous networks responses – distinct pathways of residues exhibiting altered ps-ns dynamics in response to perturbations. This type of response suggests that dynamics can play a role in the mechanical/sequential propagation of signals and that NMR studies can provide insights into specific mechanisms of allosteric propagation. Other studies show that protein dynamics can alter protein properties through the relationship with conformational entropy. These thermodynamics-based analyses show that changes in fast protein dynamics impact the energetics of the protein conformational ensemble and thus might be harnessed to select certain states with particular properties. These lessons will be useful in efforts to engineer protein function *de novo* and to optimize known protein/ligand interactions.

Relaxation-Dispersion NMR Reveals the Kinetics and Thermodynamics of Population Interconversion

The ensemble view of intra-protein communication holds that proteins exist as a continuum of interconverting structures, with the population of any particular conformer determined by the Boltzmann equilibrium. If two (or several) low energy structures have small or moderate energetic interconversion barriers between them, then they will all be populated to some degree. Using NMR relaxation-dispersion methodology, it is possible under appropriate conditions to experimentally observe these minor conformers and to determine the thermodynamics and kinetics of interconversion (for specifics concerning these methods, see [46,47,21]). As discussed in previous sections, dynamics on the ps-ns time scale are studied by measuring, among other things, the transverse relaxation rate R_2 . If a certain residue also experiences slower motions on the μ s-ms time scale, the measured rate $R_{2,obs}$ will actually be the sum of the intrinsic rate $R_{2,0}$ and a contribution from the slower dynamics known as R_{ex} . If a particular NMR-active nucleus is in fast exchange with a second conformation, then the R_{ex} contribution is proportional to $p_A p_B (\Delta\omega^2)/k_{ex}$, where p_A and p_B refer respectively to the populations of states A and B, $\Delta\omega$ is the difference between the chemical shifts of the two conformations for the nucleus in question, and k_{ex} is the rate of interconversion (the sum of the forward and reverse rates) between the two states. By measuring the transverse relaxation rate $R_{2,obs}$ as a function of the delay between refocusing elements in the NMR pulse sequence, the contribution from R_{ex} can be determined, and measurements of R_{ex} at multiple static magnetic field strengths allow some combination of $\Delta\omega$, k_{ex} , and the populations to be fit robustly. Note that the populations p_A and p_B give information on the position of the equilibrium between the two states and thus the free energy of interconversion, while k_{ex} reports on the kinetics of interconversion.

Numerous recent studies employing the relaxation-dispersion methodology have revealed connections between slow μ s-ms dynamics and protein function. In the prolyl cis-trans isomerase cyclophilin A, relaxation-dispersion experiments detected conformational switching between a major and a minor conformation of the enzyme in its free state, a key facet of the MWC/ensemble view of intra-protein communication; more interestingly, the global k_{ex} value determined for the conformational transition was found to be very similar to the sum of the rate constants for proline cis→trans and trans→cis isomerization, leading the authors to speculate that the dynamic interconversion between the two conformational states may be the limiting step in overall turnover by cyclophilin A [48].

Another elegant NMR study suggests that changes in internal dynamics over the ps-ns and μ s-ms time scales are the fundamental source of the negative cooperativity of cAMP binding to the dimeric catabolite activator protein (CAP). By characterizing the internal dynamics of CAP in the free state and with one and both CAP monomers bound to cAMP, researchers showed that binding of one cAMP molecule has little effect on the backbone dynamics throughout the dimer and even promotes increased motion on the μ s-ms time scale. However, binding of cAMP to the second CAP monomer results in large increases in ps-ns rigidity throughout the dimer as well as the quenching of most μ s-ms conformational exchange processes, which must incur a sizeable conformational entropic penalty. Based on additional NMR evidence that the structure of the dimer does not change upon binding the first cAMP, the authors suggest that the unfavorable change in conformational entropy due to increased rigidity on the ps-ns and μ s-ms time scales upon binding of the second cAMP is the fundamental source of the negative cooperativity of the cAMP/CAP system [49].

In contrast to the cAMP/CAP system, where quenched motions resulted in negative cooperativity, quenching μ s-ms motions in the Runt/CBF β heterodimer in core binding factors (CBF) appears to promote the Runt domain's DNA binding activity [50]. Testing the theory that the binding of CBF β to Runt enhances Runt's affinity for DNA by locking regions of high

conformational plasticity in Runt into a productive binding conformation, the authors screened for a single mutant showing enhanced DNA binding in the absence of CBF β and then characterized the dynamics of the mutant to see whether the enhanced DNA binding was accompanied by the same quenching of motions normally seen upon binding to CBF β . They identified the mutant R164N as having a stronger affinity for DNA, and NMR measurements revealed that this mutation resulted in greatly reduced conformational exchange in regions known to be quite flexible in WT Runt. A final test revealed that the binding affinity of the R164N mutant was much less susceptible to enhancement by CBF β binding. Taken collectively, the results suggest that the positive allosteric effect of CBF β binding to Runt is modulated by the quenching of μ s-ms motions in Runt and that these effects can be triggered by other perturbations to Runt such as mutation.

Overall, relaxation-dispersion NMR experiments are a powerful method for revealing the characteristics of conformational substates that proteins sample due to their ensemble character. These studies suggest the possibility of tuning protein function by introducing mutations or other perturbations that affect either the thermodynamic (populations) or kinetic (rate of interconversion) equilibria between the two (or more) substates. In fact, it is known that mutations in certain enzymes alter or “decouple” the μ s-ms time scale dynamics observed in the absence of mutation and thus impact enzymatic function [48,51,52]. These observations are all consistent with an ensemble-based model of allostery in which perturbing a system results in a redistribution of the most probable states of that system, leading to the new highest probability states having either an enhanced or reduced functional competency.

NMR Insights in Correlated Motions

The previous parts of Section II have presented many examples of the power of NMR spectroscopy to characterize internal protein dynamics occurring over various time scales and to detect specific changes in these dynamics in response to various perturbations. The collected evidence indicates that internal dynamics can effect long-range intra-protein communication either by modifying conformational entropy or by mechanically linking distal sites through the propagation of dynamical changes in response to perturbation. A particularly important open question in the field of protein dynamics is to what degree the internal motions of proteins might be correlated and what influence correlated motions might have on functional properties. NMR order parameters, which quantify motion on the sub-nanosecond time scale, can be used to estimate the conformational entropy of a system [53], but this method cannot account for correlated motions, and thus conformational entropy calculated using ps-ns S^2 values may be somewhat overestimated. To examine the relationship between correlated motions and conformational entropy, Stone and coworkers performed a covariation analysis of the ps-ns NMR dynamics parameters for a set of 10 mutants of the B1 domain from protein G. Their analysis revealed a long-range network comprised of residues sensitive to the dynamics of the other members of the network. The correlations among the dynamics of individual members of the network ranged from only slight to modest, however, so the authors conclude that conformational entropies calculated on the basis of NMR order parameters are probably overestimated only to a small degree [54]. A similar conclusion was reached using computational simulations [55].

Protein dynamics can also be investigated by NMR methods other than traditional relaxation experiments. Residual dipolar couplings (RDCs) are sensitive to motions occurring across a vast time scale ranging from seconds to faster than nanoseconds. By exploiting the inherent information about dynamics contained in an extensive set of RDC and hydrogen bond scalar coupling measurements of protein G, Blackledge and coworkers have detected a series of slow correlated motions at alternating positions along the β -strands comprising protein G's β -sheet [56]. Furthermore, the motions across individual β -strands appear to be mediated by the

hydrogen bonds linking the strands. The authors conclude the analysis by noting that the largest-amplitude motions in the β -sheet coincide with those residues that interact with protein G's biological partner and speculate that the increased conformational sampling at those positions may facilitate the intermolecular association process.

III. An Interdisciplinary Field: Novel Methods for Studying Intra-Protein Communication

The purpose of the final section of this review article is to provide the biological scientist with a general introduction to the application of network and information theoretic methods to the study of individual biological macromolecules. Although the application of information and network theoretic approaches to biological questions is not as developed as most traditional biological and biophysical techniques, this article will present the contributions made by such approaches thus far and highlight how they might be further developed in the future.

Protein Structural Networks and the Associated Terminology

Protein structural networks are in essence a map of the topology of a given protein with reference to a particular characteristic. Construction of a protein graph allows the protein to be analyzed using concepts from network theory without regard to the specific identities of the chemical units forming the real protein structure. The protein graph that is easiest to comprehend is the graph of protein van der Waals interactions. Every amino acid residue in the protein is considered to be a "node" or "vertex," and the physical connectivity of the entire protein can be visualized by connecting each node to every other node with which it is in van der Waals contact by drawing an "edge" between them. Note that the existence of an edge between any two nodes depends only on their spatial relationship in three dimensions and has no relation to distance in the primary sequence or the physical characteristics of the particular amino acids. In this manner, the specific characteristics of the protein have been replaced by an abstract graph based only on spatial relationships. Numerous mathematical descriptors exist for characterizing network properties (for more information, see [57]), but only those commonly used and discussed for biological systems will be presented here. The *degree* of a given node is simply the number of other nodes in the graph to which it is connected. The path length between any two nodes in a graph is defined as the minimum number of edges that must be traversed to reach the second node when starting at the first. The *average path length* of a given node is the average of the path lengths between the given node and every other node in the protein. The higher this length, the more isolated a node is from other nodes. The *clustering coefficient* of a given node measures that node's local connectivity by characterizing how many neighbors of a particular node are also neighbors of each other. The final concept important for discussing protein graphs is a node's *betweenness or betweenness centrality*, which for a given node i is defined as the number of pairs of other nodes (j,k) for which i is located on the shortest path between j and k .

Graphs can be classified into various categories according to the properties listed above. A random graph is one in which all vertices are connected randomly to other vertices. As a consequence, nodes in random graphs have small average path lengths and small clustering coefficients. At the other extreme are regular graphs, in which few long distance connectivities exist, resulting in high average path lengths and high local clustering. It is between these two extremes where protein structural graphs are found. Studies have shown that protein structural graphs generally form *small world networks*, networks that are characterized by relatively small average path lengths but relatively high clustering [58,59]. This means that proteins form networks in which most of a node's edges lead to neighbors nearby in primary sequence, but that a node will occasionally have long-distance connections to residues far away in primary

sequence. It is from the small world point of view that graph theory has been applied to questions in protein science.

The network properties of protein monomers and changes in these properties upon dimerization have been studied extensively from the small world point of view. Del Sol and coworkers have shown that dimerization involves the merging of each monomer's individual small world network into a small world network covering the whole dimer [60]. They further determined that, while the distributions of degree and betweenness were the same in the monomeric and dimeric small world networks, the location of the nodes with high betweenness changed. Nodes with high betweenness in the dimer appear at the dimerization interface with statistically significant frequency (although some nodes with high betweenness in the complex also have high betweenness in the monomeric state [61]). Most interestingly, the authors noted a connection between the network property of betweenness and a real-world thermodynamic observable, namely the free energy contributed by individual residues to the binding process. Those dimerization interface residues identified by scanning mutagenesis as binding "hotspots" were the same as nodes of high betweenness to a significant degree of the time. This tendency reveals the ability of network analysis to examine and relate to physically observable phenomena. In fact, network analyses have also contributed to problems as varied as predicting catalytic site residues in enzymes [62] and predicting residues comprising the folding nucleus in protein folding reactions [63–65,58].

With regard to intra-protein communication, it is tempting to view properties like high betweenness centrality and low average path length as possible identifiers of residues that are members of pathways of long distance communication. Indeed, there is some evidence that this might be the case in chymotrypsin inhibitor 2 (CI2) [59]. The chief drawback hindering network analysis from making a greater contribution to understanding dynamic processes in proteins is its dependence on static structures for the generation of the initial graph upon which all subsequent analysis is based. Such static graphs do not take into account the multitude of conformational fluctuations and their role in protein function that have been the focus of this review. Efforts are underway, however, to extend network analysis in protein science by basing graphs on features other than the average 3D structure. Recent attempts have, for example, constructed graphs with edges weighted by the contact energy between nodes rather than simply the existence of a contact; this has the advantage of giving a subset of a node's contacts a higher importance than the rest. The graph weighted by contact energy was used to study the CI2 unfolding reaction from the point of view of changing network characteristics [66]. Using evolutionary data in concert with the graph theoretic approach, Schmidt and coworkers were able to identify conserved putative pathways of allosteric communication between the ATP binding site and lever arm in myosin. By aligning numerous myosin sequences from multiple species and then constructing a directed structural graph with edges weighted according to the degree of evolutionary conservation of the destination vertex, the researchers were able to suggest a pathway conserved across multiple species that physically couples ATP hydrolysis to the lever arm's recovery stroke [67]. This study is notable because it is one of the few thus far to combine graph theoretic approaches with other types of data available to biological scientists. Other applications of protein structural graphs have sought additional insight by considering the properties of subgraphs of the main contact graph, such as the respective subnetworks formed by hydrophobic, hydrophilic, and charged amino acids [68]. Most needed, however, is a method for constructing graphs that take into account flexibility and the dynamical nature of proteins; these graphs should have a better likelihood of providing meaningful insight into mechanisms of intra-protein communication.

IV. Conclusions

In this article, we have attempted to provide an overview of the development and current standing of frameworks for understanding long-range communication in proteins, particularly with regard to cooperativity and allostery. Both the mechanical/sequential model and the thermodynamic ensemble model have been successful in explaining cooperative interactions in a wide variety of systems, but questions remain as to the general applicability of each. The mechanical/sequential model, for example, has difficulty explaining energetic couplings between residues that do not lie on the putative communication pathway, while the thermodynamic ensemble model does not necessarily explain why distinct communication pathways appear to exist for some proteins and not others. It will be in the unification or generalization of these ideas where the next advance in understanding intra-protein communication occurs. We have also discussed two major methods for investigating these issues. NMR spectroscopy is a powerful method for experimentally probing the dynamical features of proteins, and graph theoretic approaches have appeared as a novel way to analyze the properties of proteins from a structural point of view. (A complete review of computational and MD simulations applied to these issues, while extremely relevant and powerful, is beyond the scope of this review.) Due to the complexity and fundamental importance of the phenomenon of intra-protein communication, we suggest that it is only through creative, multidisciplinary approaches that ultimate insight will be achieved, and we hope that this review will contribute to a greater understanding of the scope of the question as well as of the tools available to study it.

Abbreviations

ATCase, aspartate transcarbamoylase
 cAMP, cyclic AMP
 CAP, catabolite activator protein
 HEWL, hen egg-white lysozyme
 KNF, Koshland-Némethy-Filmer or sequential model
 Kv, voltage-activated potassium channel
 MD, molecular dynamics
 MWC, Monod-Wyman-Changeux or concerted model
 NMR, nuclear magnetic resonance
 NOE, nuclear Overhauser enhancement
 RDCs, residual dipolar couplings
 R_{ex} , line-broadening contribution to R_2 from μs -ms time scale motions
 S^2 , generalized order parameter for motions on the ps-ns time scale
 SAXS, small angle x-ray scattering
 T_1 , spin-lattice relaxation time constant (equals $1/R_1$)
 T_2 , spin-spin relaxation time constant (equals $1/R_2$)
 τ_c , correlation time for internal motion on the ps-ns time scale
 WT, wild type

Acknowledgements

This work was supported by an NSF Graduate Research Fellowship (to M.J.W.), NIH Grant GM066009 (to A.L.L.), and NSF Grant MCB-0344354 (to A.L.L.)

References

1. Koshland DE. Proc. Natl. Acad. Sci. USA 1958;44:98–104. [PubMed: 16590179]
2. Koshland DE, Nemethy G, Filmer D. Biochemistry 1966;5:365–385. [PubMed: 5938952]

3. Purohit P, Mitra A, Auerbach A. *Nature* 2007;446:930–933. [PubMed: 17443187]
4. Yifrach O, MacKinnon R. *Cell* 2002;111:231–239. [PubMed: 12408867]
5. Horovitz A, Fersht AR. *J. Mol. Biol* 1990;214:613–617. [PubMed: 2388258]
6. Hidalgo P, MacKinnon R. *Science* 1995;268:307–310. [PubMed: 7716527]
7. Sadovalsky E, Yifrach O. *Proc. Natl. Acad. Sci. USA* 2007;104:19813–19818. [PubMed: 18077413]
8. Monod J, Wyman J, Changeux J-P. *J. Mol. Biol* 1965;12:88–118. [PubMed: 14343300]
9. Fetler L, Kantrowitz ER, Vachette P. *Proc. Natl. Acad. Sci. USA* 2007;104:495–500. [PubMed: 17202260]
10. Velyvis A, Yang YR, Schachman HK, Kay LE. *Proc. Natl. Acad. Sci. USA* 2007;104:8815–8820. [PubMed: 17502625]
11. Hilser VJ, García-Moreno EB, Oas TG, Kapp G, Whitten ST. *Chem. Rev* 2006;106:1545–1558. [PubMed: 16683744]
12. Freire E. *Proc. Natl. Acad. Sci. USA* 1999;96:10118–10122. [PubMed: 10468572]
13. Koshland DE Jr. *Nat. Med* 1998;4:1112–1114. [PubMed: 9771734]
14. Hilser VJ, Townsend BD, Freire E. *Biophys. Chem* 1997;64:69–79. [PubMed: 9127939]
15. Hilser VJ, Dowdy D, Oas TG, Freire E. *Proc. Natl. Acad. Sci. USA* 1998;95:9903–9908. [PubMed: 9707573]
16. Pan H, Lee JC, Hilser VJ. *Proc. Natl. Acad. Sci. USA* 2000;97:12020–12025. [PubMed: 11035796]
17. Formanek MS, Ma L, Cui Q. *Proteins* 2006;63:846–867. [PubMed: 16475196]
18. Ma L, Cui Q. *J. Am. Chem. Soc* 2007;129:10261–10268. [PubMed: 17655236]
19. Cooper A, Dryden DTF. *Eur. Biophys. J* 1984;11:103–109. [PubMed: 6544679]
20. Wand AJ. *Nat. Struct. Biol* 2001;8:926–931. [PubMed: 11685236]
21. Kay LE. *J. Magn. Reson* 2005;173:193–207. [PubMed: 15780912]
22. Jarymowycz VA, Stone MJ. *Chem. Rev* 2006;106:1624–1671. [PubMed: 16683748]
23. Lipari G, Szabo A. *J. Am. Chem. Soc* 1982;104:4546–4559.
24. Lipari G, Szabo A. *J. Am. Chem. Soc* 1982;104:4559–4570.
25. Muhandiram DR, Yamazaki T, Sykes BD, Kay LE. *J. Am. Chem. Soc* 1995;117:11536–11544.
26. Millet O, Muhandiram DR, Skrynnikov NR, Kay LE. *J. Am. Chem. Soc* 2002;124:6439–6448. [PubMed: 12033875]
27. Igumenova TI, Frederick KK, Wand AJ. *Chem. Rev* 2006;106:1672–1699. [PubMed: 16683749]
28. Hu H, Clarkson MW, Hermans J, Lee AL. *Biochemistry* 2003;42:13856–13868. [PubMed: 14636053]
29. Clarkson MW, Lee AL. *Biochemistry* 2004;43:12448–12458. [PubMed: 15449934]
30. Clarkson MW, Gilmore SA, Edgell MH, Lee AL. *Biochemistry* 2006;45:7693–7699. [PubMed: 16784220]
31. Gunasekaran K, Ma B, Nussinov R. *Proteins* 2004;57:433–443. [PubMed: 15382234]
32. Igumenova TI, Lee AL, Wand AJ. *Biochemistry* 2005;44:12627–12639. [PubMed: 16171378]
33. Millet O, Mittermaier A, Baker D, Kay LE. *J. Mol. Biol* 2003;329:551–563. [PubMed: 12767834]
34. Mittermaier A, Kay LE. *Protein Sci* 2004;13:1088–1099. [PubMed: 15044737]
35. Johnson E, Chazin WJ, Rance M. *J. Mol. Biol* 2006;357:1237–1252. [PubMed: 16476440]
36. Zhuravleva A, Korzhnev DM, Nolde SB, Kay LE, Arseniev AS, Billeter M, Orekhov VY. *J. Mol. Biol* 2007;367:1079–1092. [PubMed: 17306298]
37. Zheng W, Brooks BR, Doniach S, Thirumalai D. *Structure* 2005;13:565–577. [PubMed: 15837195]
38. Fuentes EJ, Der CJ, Lee AL. *J. Mol. Biol* 2004;335:1105–1115. [PubMed: 14698303]
39. Lockless SW, Ranganathan R. *Science* 1999;286:295–299. [PubMed: 10514373]
40. Ota N, Agard DA. *J. Mol. Biol* 2005;351:345–354. [PubMed: 16005893]
41. Sharp K, Skinner JJ. *Proteins* 2006;65:347–361. [PubMed: 16933296]
42. Fuentes EJ, Gilmore SA, Mauldin RV, Lee AL. *J. Mol. Biol* 2006;364:337–351. [PubMed: 17011581]
43. Namanja AT, Peng T, Zintsmaster JS, Elson AC, Shakour MG, Peng JW. *Structure* 2007;15:313–327. [PubMed: 17355867]

44. Peng T, Zintsmaster JS, Namanja AT, Peng JW. *Nat. Struct. Mol. Biol* 2007;14:325–331. [PubMed: 17334375]
45. Frederick KK, Marlow MS, Valentine KG, Wand AJ. *Nature* 2007;448:325–329. [PubMed: 17637663]
46. Palmer AG 3rd. *Curr. Opin. Struct. Biol* 1997;7:732–737. [PubMed: 9345634]
47. Palmer AG 3rd, Kroenke CD, Loria JP. *Methods Enzymol* 2001;339:204–238. [PubMed: 11462813]
48. Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D. *Nature* 2005;438:117–121. [PubMed: 16267559]
49. Popovych N, Sun S, Ebright RH, Kalodimos CG. *Nat. Struct. Mol. Biol* 2006;13:831–838. [PubMed: 16906160]
50. Li Z, Lukasik SM, Liu Y, Grembecka J, Bielnicka I, Bushweller JH, Speck NA. *J. Mol. Biol* 2006;364:1073–1083. [PubMed: 17059830]
51. Kovrigina EL, Loria JP. *Biochemistry* 2006;45:2636–2647. [PubMed: 16489757]
52. Watt ED, Shimada H, Kovrigina EL, Loria JP. *Proc. Natl. Acad. Sci. USA* 2007;104:11981–11986. [PubMed: 17615241]
53. Yang D, Kay LE. *J. Mol. Biol* 1996;263:369–382. [PubMed: 8913313]
54. Mayer KL, Earley MR, Gupta S, Pichumani K, Regan L, Stone MJ. *Nat. Struct. Biol* 2003;10:962–965. [PubMed: 14528292]
55. Prompers JJ, Bruschweiler R. *J. Phys. Chem. B* 2000;104:11416–11424.
56. Bouvignies G, Bernado P, Meier S, Cho K, Grzesiek S, Bruschweiler R, Blackledge M. *Proc. Natl. Acad. Sci. USA* 2005;102:13885–13890. [PubMed: 16172390]
57. Bonchev D, Buck GA. *J. Chem. Inf. Model* 2007;47:909–917. [PubMed: 17407281]
58. Vendruscolo M, Dokholyan NV, Paci E, Karplus M. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys* 2002;65:061910. [PubMed: 12188762]
59. Atilgan AR, Akan P, Baysal C. *Biophys. J* 2004;86:85–91. [PubMed: 14695252]
60. del Sol A, Fujihashi H, O'Meara P. *Bioinformatics* 2005;21:1311–1315. [PubMed: 15659419]
61. del Sol A, O'Meara P. *Proteins* 2005;58:672–682. [PubMed: 15617065]
62. Chea E, Livesay D. *BMC Bioinformatics* 2007;8:153. [PubMed: 17498304]
63. Shakhnovich E, Abkevich V, Ptitsyn O. *Nature* 1996;379:96–98. [PubMed: 8538750]
64. Vendruscolo M, Paci E, Dobson CM, Karplus M. *Nature* 2001;409:641–645. [PubMed: 11214326]
65. Dokholyan NV, Li L, Ding F, Shakhnovich EI. *Proc. Natl. Acad. Sci. USA* 2002;99:8637–8641. [PubMed: 12084924]
66. Jiao X, Chang S, Li CH, Chen WZ, Wang CX. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys* 2007;75:051903. [PubMed: 17677094]
67. Tang S, Liao JC, Dunn AR, Altman RB, Spudich JA, Schmidt JP. *J. Mol. Biol* 2007;373:1361–1373. [PubMed: 17900617]
68. Aftabuddin M, Kundu S. *Biophys. J* 2007;93:225–231. [PubMed: 17172302]

Fig. 1.

A schematic summary of the two basic frameworks for understanding long-range intra-protein communication. (a) The sequential (KNF) model whereby a series of small conformational changes resulting from induced fit ligand binding propagates to a distal site in the protein, changing its conformation and making the binding of a second ligand more favorable. (b) The conformational ensemble viewpoint developed from the MWC model of ligand binding. The high and low affinity states are populated at all times, though to differing degrees. Ligand binding by the high affinity state, although relatively rare on account of its smaller population, stabilizes that state and redistributes the populations in the conformational ensemble, leading to the eventual dominance of the high affinity state.

Fig. 2.

A representation of the conformational ensemble framework for long-range communication. Perturbing the system can lead to two possible outcomes. (a) Perturbation shifts the original population distribution (gray curve) to a new distribution having a global energetic minimum at a new average conformation (black curve). (b) The ensemble is perturbed such that the global minimum (and thus the average conformation) of the original distribution (gray curve) does not change; rather, the width of the distribution becomes either wider (dotted curve) or narrower (black curve), representing a respective gain or loss of conformational entropy.

Fig. 3.

A stereoscopic view of the consequences of peptide binding to PDZ2 from human tyrosine phosphatase 1E (adapted from PDB code 1D5G). The bound peptide (black) perturbs the ps-ns dynamics of methyl-bearing side chains distal to the binding site. The perturbed distal side chains are shown as gray van der Waals surfaces. For clarity, perturbed side chains that directly interact with the peptide are not shown.