The role of protein conformational fluctuations in allostery, function, and evolution

James O. Wrabl [a,b], Jenny Gu [a,1], Tong Liu [a,2], Travis P. Schrank [a], Steven T. Whitten [a,3], Vincent J. Hilser [a,b,*]

[a] Departments of Biology and Biophysics, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218 USA
[b] Department of Biochemistry and Molecular Biology, University of Texas Medical Branch at Galveston, 301 University Boulevard, Galveston, TX 77555-1068 USA

ABSTRACT

It is now well-known that proteins exist at equilibrium as ensembles of conformational states rather than as unique static structures. Here we review from an ensemble perspective important biological effects of such spontaneous fluctuations on protein allostery, function, and evolution. However, rather than present a thorough literature review on each subject, we focus instead on connecting these phenomena through the ensemble-based experimental, theoretical, and computational investigations from our laboratory over the past decade. Special emphasis is given to insights that run counter to some of the prevailing ideas that have emerged over the past 40 years of structural biology research. For instance, when proteins are viewed as conformational ensembles rather than as single structures, the commonly held notion of an allosteric pathway as an obligate series of individual structural distortions loses its meaning. Instead, allostery can result from energetic linkage between distal sites as one Boltzmann distribution of states transitions to another. Additionally, the emerging principles from this ensemble view of proteins have proven surprisingly useful in describing the role of intrinsic disorder in inter-domain communication, functional adaptation mediated by mutational control of fluctuations, and evolutionary conservation of the energetics of protein stability.

© 2011 Published by Elsevier B.V.

1. Introduction

Proteins are essential to life because of the manifold of functional roles they perform in organisms [1–3]. One characteristic that allows proteins to adopt such diverse functions is structural malleability [4–7]. It is thus simultaneously correct to describe a particular protein as having a "folded" or "ground state" structure as well as undergoing "spontaneous fluctuations" or transitions to "excited states" around that structure. By definition, information from these conformational fluctuations cannot be obtained directly from inspection of a single ground state structure[8,9]. In spite of this limitation, much experimental evidence exists that underscores the functional importance of excited as well as ground states [10–16].

Conformational differences between ground and excited states are caused by local variations in energy. Such variation may be sourced from ambient thermal motion, kT, or from a binding or catalytic event. Regardless of the physical source of the energy, the effect on protein conformation will depend on multiple factors, including (but not limited to) enthalpic protein–protein and protein–solvent interactions, as well as the entropic change to the protein–solvent system associated with populating the excited state. Clearly, a complete structural and functional understanding of a particular protein can only come from an understanding of the equilibria involving these structural perturbations. Unfortunately, however, knowledge of the ground state of the protein (in the form of a high-resolution crystal structure) only provides part of that equilibrium, leaving the excited states as largely unknown.

How can detailed knowledge about the energetics of a protein and its effect on spontaneous conformational fluctuations be obtained? Can this knowledge be used to rationalize protein allostery? Since the pioneering work of Monod–Wyman–Changeux [17] and Koshland–Nemethy–Filmer [18], many investigations have addressed these fundamental questions theoretically, experimentally and computationally. Examples of the variety of approaches include double-mutant cycles [19], correlated mutations analysis [20], changes in dynamics [10,21–23] and population shifts [24,25].

The approach of this laboratory, partly inspired by fruitful discussions during the early Gibbs Conferences on Biothermodynamics, was to simply model a protein as an ensemble of conformations at equilibrium [26–29]. The unique contribution of the model was to provide conceptual access to the energetics of the excited states of the ensemble.

* Corresponding author at: Johns Hopkins University, Department of Biology, Mudd Hall Room 117A, 3400 N. Charles Street, Baltimore, MD 21218. Tel.: +1 410 516 6072; fax: +1 410 516 5213.
E-mail addresses: jowrabl@jhu.edu (J.O. Wrabl), j.gu@uni-muenster.de (J. Gu), tongliu@ucsd.edu (T. Liu), tpschran@utmb.edu (T.P. Schrank), steve.whitten@txstate.edu (S.T. Whitten), hilser@jhu.edu (V.J. Hilser).
[1] Present address: Evolutionary Bioinformatics Group, Institute of Evolution and Biodiversity, Westfaelische Wilhelms University, Hufferstrasse 1, D48149 Muenster, Germany.
[2] Present address: Department of Medicine and Biomedical Sciences, University of California San Diego, La Jolla, CA 92093-0656, USA.
[3] Present address: Department of Chemistry and Biochemistry, Texas State University, San Marcos, 601 University Drive, San Marcos, TX 78666, USA.

Thus, estimation of the complete equilibrium between ground and excited states, the conformational fluctuations of the protein, and the global and local stability of the protein all became possible, albeit at a seemingly low level of energetic detail as compared to other approaches [22]. As discussed in the following biased retrospective, detailed knowledge of the protein ensemble has proven effective and invaluable in providing insight into many complex biological phenomena, including allostery and functional adaptation [30–38]. The biological importance of the ensemble description of proteins is further underscored by the emerging observation that spontaneous fluctuations of proteins are largely conserved during evolution [39–42].

2. Modeling the energetic and conformational diversity of the protein ensemble

Since the ground state of a protein is a necessary component of spontaneous fluctuations, this laboratory's model [26,29,36,43] of the native state ensemble starts with the experimentally observed crystal structure input as a reference state. A key assumption of the model is that the crystal structure represents the true ground state of the protein. Generation of the ensemble, through an algorithmic process named COREX/BEST [43] (http://best.bio.jhu.edu; Fig. 1), proceeds by systematically unfolding all regions of the structure in units of a predefined window size of amino acid residues. Importantly, this process is exhaustive—every possible permutation of folded and unfolded units is sampled, from all of the protein's amino acids being in the crystal structure conformation to all of the amino acids being unfolded.

The COREX/BEST approach to conformational sampling is akin to the coarse-grained approach developed in helix-coil theory [44], where a residue has two possible conformational states. In the latter approach, these states are helix (h) or coil (c), thus a molecular conformation is a string of "h's" and "c's" and there are 2^N microstates that comprise the partition function for a protein chain of length N. The COREX/BEST approach is similar, the difference being that a residue may be either folded (F) or unfolded (U), and the 2^N microstates are made more computationally tractable through the introduction of a sliding window, usually five residues in size. Within this window all residues are treated as either completely folded or completely unfolded. In this way, every possible conformational fluctuation is sampled, subject to the constraints of window size.

An energy function is used to compute the Boltzmann weight for each microstate i and these weights are used to calculate the partition function Q and the population of each microstate in the ensemble. The energy function, more precisely a mean field approximation of a free energy functional, is based on two physical observables: solvent accessible surface area and conformational entropy. [26] The former quantity has been experimentally parameterized from calorimetric measurements of protein denaturation [27,28,45–47], and the latter quantity has been computationally estimated from molecular dynamics simulation [48].

With the partition function in hand, many thermodynamically descriptive quantities (including free energy and its component enthalpy and entropy) can be computed [49,50]. Of fundamental interest is the stability constant, κ_f, representing the relative stability
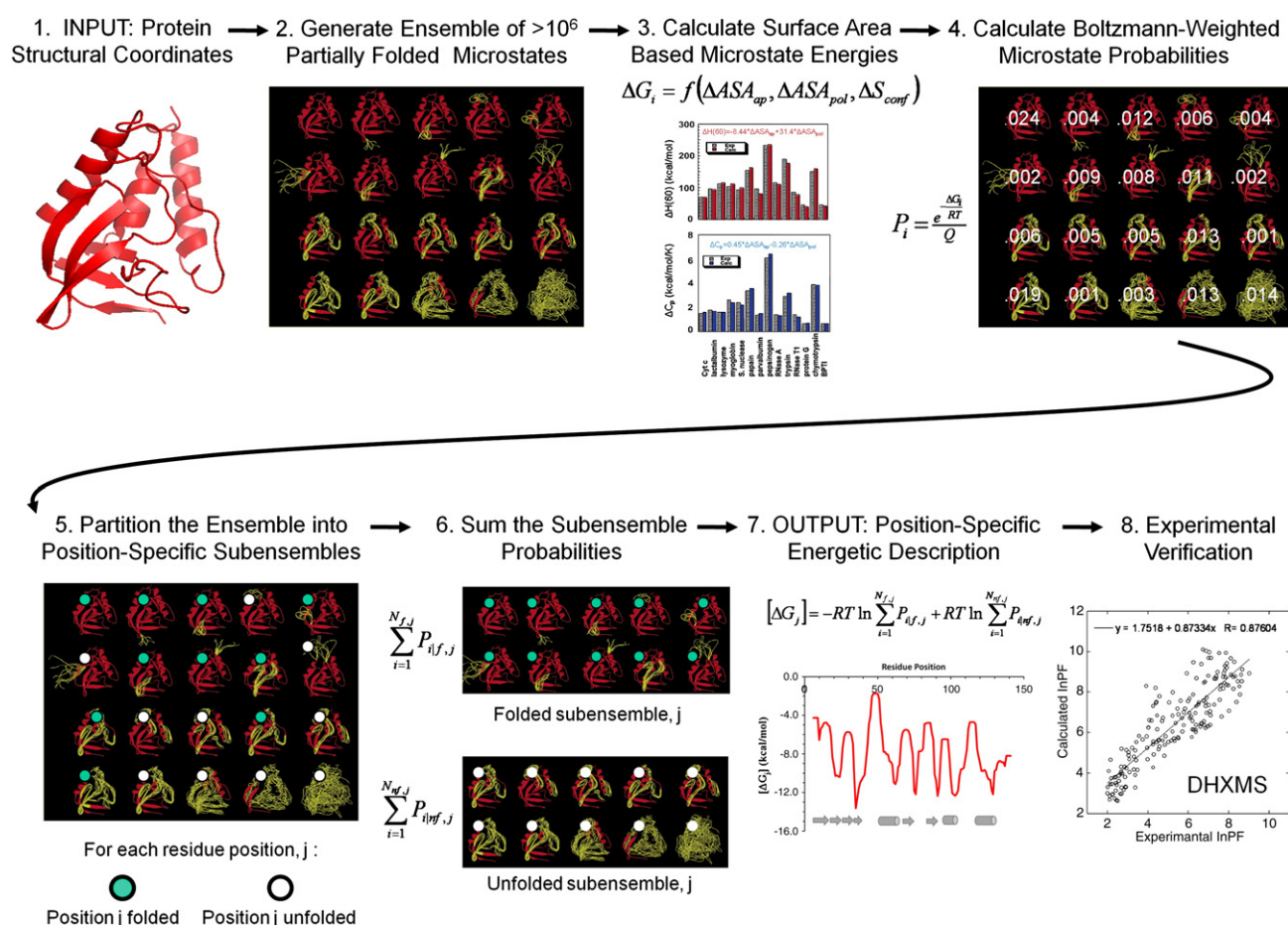

Fig. 1. Equilibrium conformational fluctuations of proteins as modeled by the ensemble-based COREX/BEST algorithm. Input to the algorithm is the experimentally-determined structural coordinates of a protein. Output is a set of ensemble-based thermodynamic descriptors for every residue in the protein, including (but not limited to) ΔG, ΔH, and TΔS. Importantly, this description has been qualitatively and quantitatively verified by experiment. The experimental verification shown in panel 8 is predicted vs. measured deuterium-hydrogen exchange mass spectrometry (DHXMS) protection factors for chicken brain alpha-spectrin 16–17 (laboratory of Virgil L. Woods, Jr., manuscript submitted for publication).

reported by each residue position j in the protein [26,50–52]. This output value is a direct measure of the conformational variability around the ground state structure: regions with high κ_f experience few conformational fluctuations at equilibrium, regions with low κ_f experience many conformational fluctuations. Crucially, the residue stability constant has been experimentally verified with hydrogen exchange protection factor measurements on many occasions (Fig. 1). This finding, taken together with agreement from other diverse experimental observables, has repeatedly confirmed that the ensemble-based approach provides a reasonable quantitative description [26,30,31,34,36,53–55].

It is important to note that local fluctuations, when viewed from the perspective of the COREX/BEST model of the ensemble, are not necessarily a consequence of only the residues involved in the fluctuations [50,52,56]. Instead, residues far from the fluctuation could potentially contribute to local fluctuations in a particular region of the protein, since the determining factor in populating a particular state is the energy of the entire microstate involved (Fig. 2). If unfolding residues distal to a particular unfolded region of the protein lowers the energy of the entire microstate, microstates unfolding both sets of residues will be preferentially populated. Correlations between
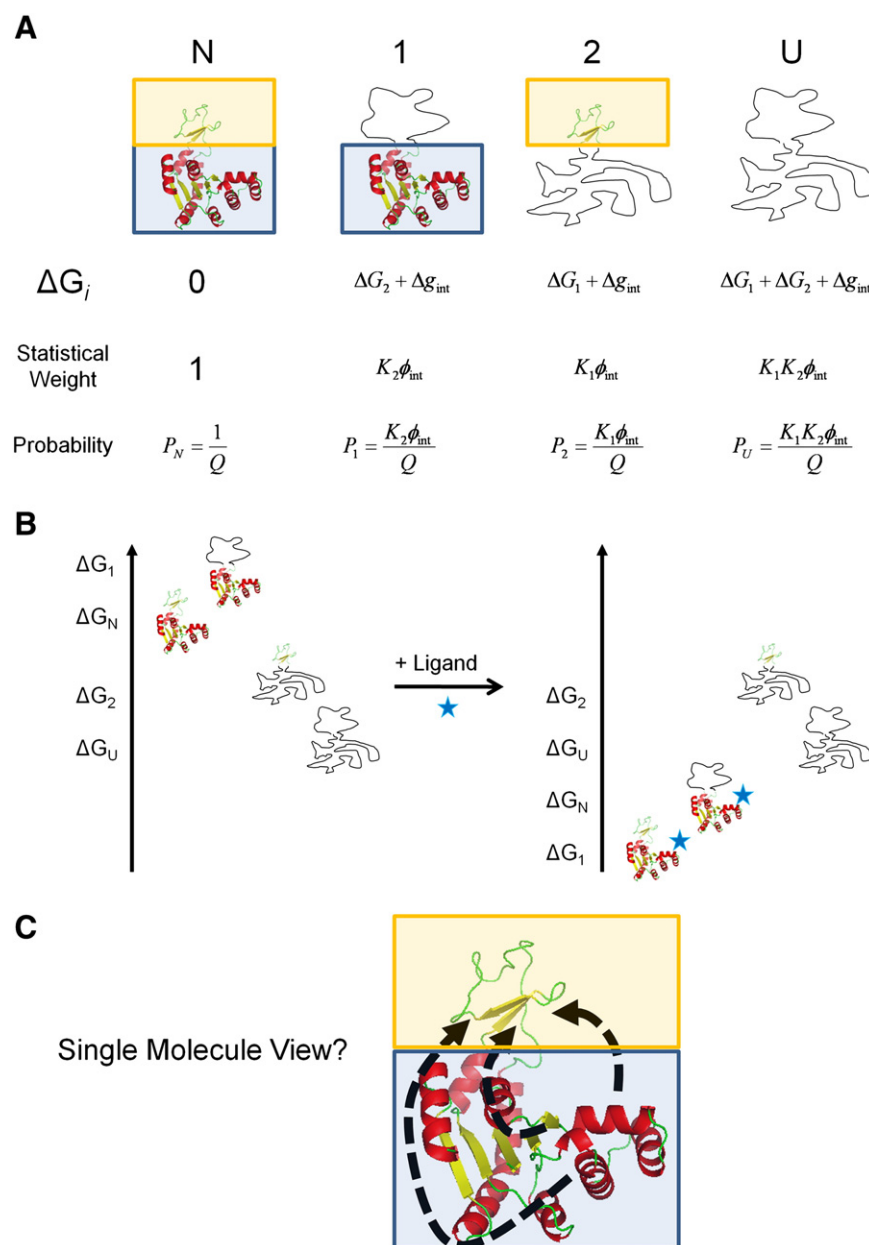

Fig. 2. Redistribution of ensemble microstates in response to perturbation mediates energy transduction within a single protein. A. A simple ensemble description of a protein containing two subdomains "1" and "2" (orange and blue boxes, respectively) and four possible microstates. Note that no structural pathway is defined between the blue and orange domains in this model. The partition function Q for this ensemble can be completely enumerated according to the equilibria and equations given in the table. B. Upon binding ligand to subdomain 2, the folded population of subdomain 1 is selectively increased from approximately 30% to 80% due to redistribution of energetically favorable microstates that bind ligand in subdomain 2; this result is experimentally observed as linkage between the two subdomains. It is again emphasized that this result obtains in the absence of a structural pathway connecting the subdomains. The biologically reasonable values used for this example are: $\Delta G_1 = -0.7$ kcal/mol, $\Delta G_2 = -2.3$ kcal/mol, $\Delta G_{interaction} = -1.6$ kcal/mol, $\Delta G_{binding} = -3.0$ kcal/mol. C. Because the individual domains or regions that make up a protein are in general cooperative, the energy cannot be assigned to any particular residue(s). In other words, result cannot be quantitatively described by physically connecting two sites in a single molecule structural picture of the protein with a schematic arrow.

fluctuations, which experimentally appear as cooperativity, are manifested in the relative probabilities of different microstates in the ensemble, without explicitly invoking a specific interaction network between cooperative residue pairs (Fig. 2). This result, known [10,21,25,57,58] but nonetheless non-intuitive to those unfamiliar with statistical mechanics, is the key to the COREX/BEST algorithm's success at capturing cooperativity between both residues in physical contact and those located in distant regions of the protein.

3. Allostery manifested as cooperative response of the ensemble to energetic perturbation

A second phenomenon that can be quantitatively described by the ensemble treatment of proteins is the redistribution of microstate populations in response to an energetic perturbation of the ensemble. This phenomenon is especially relevant biologically as it pertains to the protein's cooperative response to an allosteric signaling event, binding of ligand (Fig. 2), or effect of a stabilizing or destabilizing point mutation. Historically, this laboratory has investigated population redistribution resulting from energetic perturbations at several levels of detail: coupling between residue pairs [35,59], coupling between regions within a single protein [37], and coupling between tethered domains (cf.Fig. 4). [33]. In each case, experimental observations were recapitulated and new insights obtained with the COREX/BEST ensemble description.

Pairwise residue energetic coupling was analyzed from the ensembles of the model proteins DHFR and eglin C [35,37,59]. In both proteins, couplings existed that defined the magnitude and extent to which mutational effects propagated away from the site of the point mutation. Positive and negative couplings, sometimes between residues quite distant in the ground state structures, were observed to be in general agreement with the experiment. For instance, one striking observation was the observed connectivity between the distal loop 63–68 of DHFR and the folate binding site (Fig. 3) [37], a result that was subsequently confirmed by NMR relaxation measurements [60]. The importance of this result is that it revealed a previously undescribed phenomenon—allosteric effects that can propagate in the absence of a pathway connecting the two sites. In earlier studies [57], Cooper and Dryden made the argument on theoretical grounds that free energy changes could be manifested exclusively in changes in the widths of a conformational distribution upon the addition of an allosteric effector—in effect, that allostery can occur without a structural change. Our results confirm that observation, and go on to show that where structural and energetic changes are observed, those allosteric effects need not form structurally contiguous pathways through the protein. In effect, this result for DHFR highlighted the notion that rather than being attributed to individual residues or pathways through the protein, allostery is more appropriately viewed as a "diffusely distributed" (i.e., all protein–protein and protein–solvent interactions that stabilize a particular state are responsible for allostery), which stems from the fact that upon binding the (often highly heterogenous) conformational ensemble is redistributed, a notion that has subsequently been described by other groups [21,58].

In other results for DHFR, functional connectivity has been experimentally observed between the NADPH and folate binding sites: binding in either site affects the affinity of DHFR for the other ligand [61]. Averaging the effects of modeled energetic couplings between the sites of point mutations and the binding sites yielded excellent agreement with the experimentally observed effects on K_m[37]. Again, these energetic couplings existed in the absence of visible pathways connecting the sites, indicating that allostery could exist without a network of structurally perturbed contiguous residues (Fig. 3).

The difficulties of interpreting distant energetic couplings in terms of a structural pathway connecting the sites were brought into focus by detailed interrogation of the DHFR ensemble [35]. This study concluded that the origin of any single observed pairwise energetic
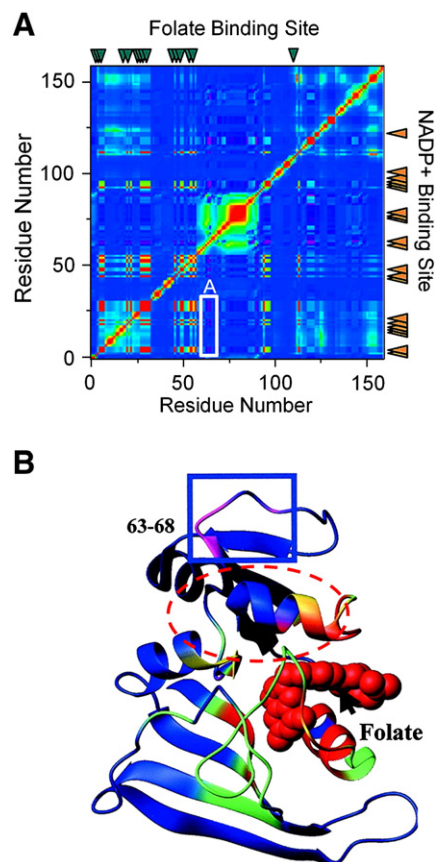

Fig. 3. One example of protein functional allostery in the absence of a structural pathway in *E. coli* dihydrofolate reductase (DHFR). A. Functional connectivities (FCs) between the folate binding site and the surrounding protein, as computed with the COREX/BEST algorithm. These FCs quantify the degree of predicted cooperative stabilization (i.e. energetic linkage) between the folate site and the rest of the protein, as previously described. [37] Red corresponds to positive linkage, blue to zero linkage, and purple to negative linkage. Predicted negative linkage between the distal loop residues 63–68 and the folate binding site (white box A) is experimentally observed. B. Energetically, the folate site is negatively coupled to distal loop 63–68 despite the absence of a connecting structural pathway: the intervening residues' energetic linkages to the folate site are essentially zero (predominance of blue color in the red dashed circle). This figure and accompanying results are adapted from Pan et al.[37]. Copyright (2000) National Academy of Sciences, USA.

coupling requires consideration of the structural and energetic properties of at least four distinct sub-ensembles. Since each sub-ensemble contained substantial structural and energetic variation within its microstates, it was not possible to rationalize energetic coupling with one (or a few) ground state conformations. Although such conclusions undermine the importance of specific interactions in transducing energy within proteins, the results also suggest that allostery is robustly encoded in the distribution.

In another example of ensemble-mediated allostery, the impact of local unfolding on the protonation state of staphylococcal nuclease (SNase) was used to demonstrate that local unfolding can be used to tune pK_a values of ionizable groups, and conversely, that pK_a values can be used to tune the pH sensitivity of local fluctuations [54]. In that study, it was shown that the pH sensitivity of the stability of SNase could be quantitatively captured (Fig. 4) with the COREX/BEST algorithm, with the ensemble being able to recapitulate the pH dependence of the proton binding (Fig. 4B). In the case of SNase, the analysis revealed that acid unfolding resulted in approximately four protons net bound by the denatured state, relative to the native state. Interestingly, the results demonstrated that the source of the cooperativity of four was not
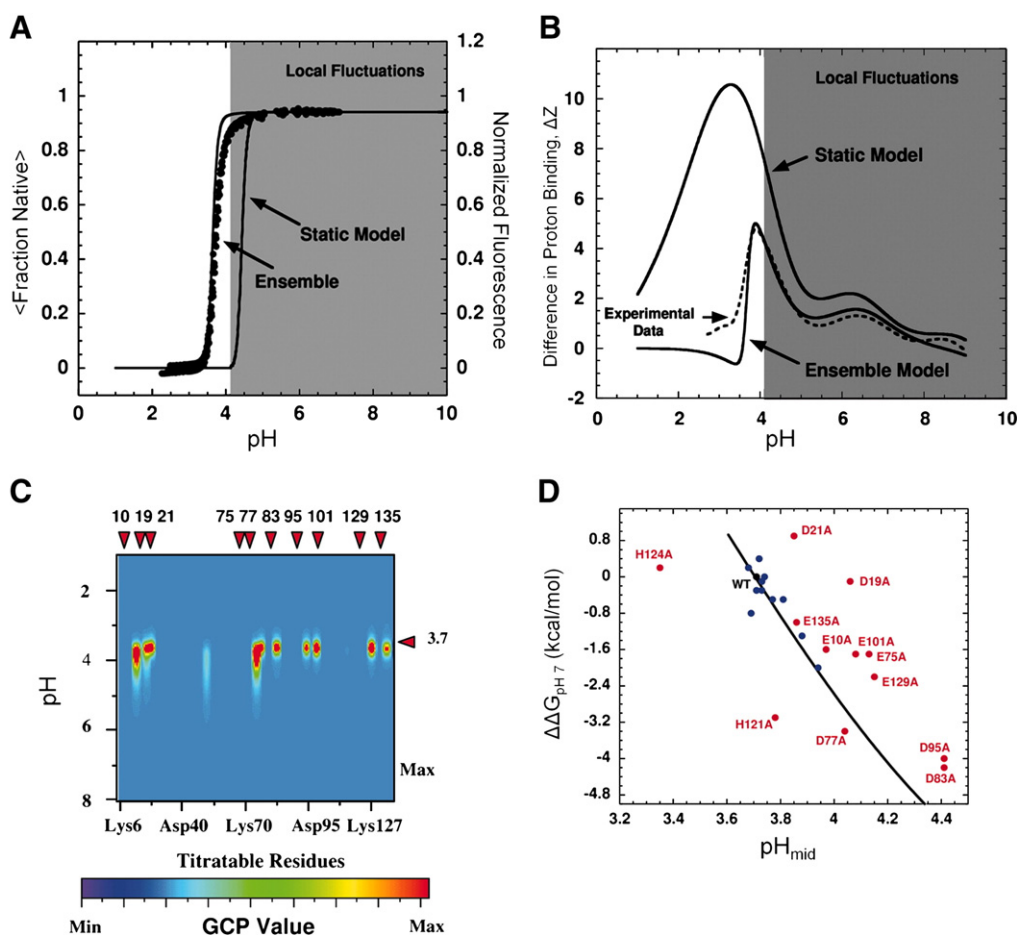
Fig. 4. A second example of protein allostery in staphylococcal nuclease (SNase) explained by ensemble modulation. A. Experimental acid-unfolding of wild-type SNase monitored by tryptophan fluorescence. A COREX/BEST ensemble-based model of the folding state of the protein describes the observed data better than a static structure model using Poisson–Boltzmann electrostatics, as previously described. [54] The gray-shaded region indicates conditions where the protein's conformational fluctuations are predominantly local. B. Experimental vs. computed proton binding in SNase. Again, the COREX/BEST ensemble-based model reproduces the observed data more faithfully than does a model based on only the fully folded and fully unfolded states of the protein, as previously described [54]. C. Global coupling parameters, GCP, were used to investigate the microscopic origins of the pH effects described in panels A and B. High GCP values (yellow and red on the accompanying scale) indicate residues whose pH titration is dependent on conformational redistribution of the ensemble. Only a subset of all ionizable residues exhibit high GCP values and are predicted to govern the energetics of the acid unfolding reaction. As previously described [54], the GCP at each residue j in the protein was defined as a function of pH: $GCP(pH)_j = (Z(pH)_{f,j} - <Z(pH)_j>) \times (Z(pH)_{nf,j} - <Z(pH)_j>)$. In this equation, the values $Z(pH)_{f,j}$ and $Z(pH)_{nf,j}$ are the pH dependent protonation states of residue j in the fully folded and fully unfolded states, respectively. $<Z(pH)_j>$ is the ensemble-weighted protonation value. D. The effects of point mutations on the stability of SNase at pH 7; $\Delta\Delta G_{pH7}$ indicates the difference in stability between the wild-type protein and the labeled mutants. The black curve describes the expected dependence; for mutants that fall on the curve (blue points), the substituted amino acid is not coupled to the global unfolding transition. Mutants that do not fall on the curve (red labeled points) are often in the subset identified in panel C, suggesting that the pH-dependent energetics of SNase folding are controlled by ensemble and not static structure effects. This figure and accompanying results are adapted from Whitten, et al.[54]. Copyright (2005) National Academy of Sciences, USA.
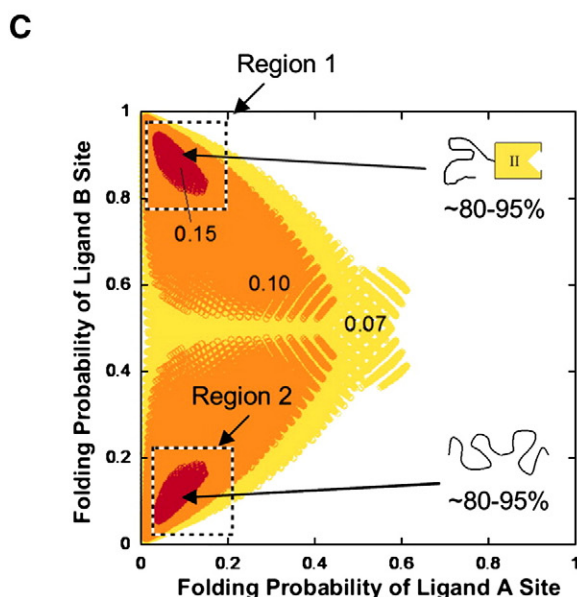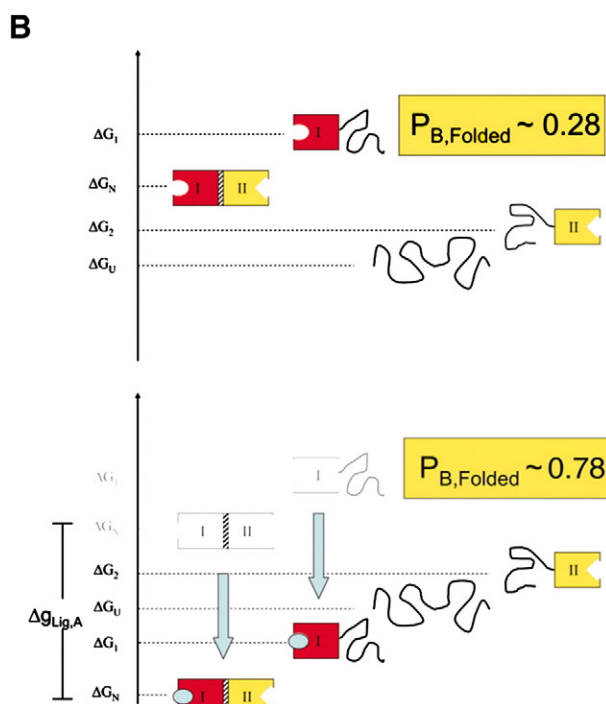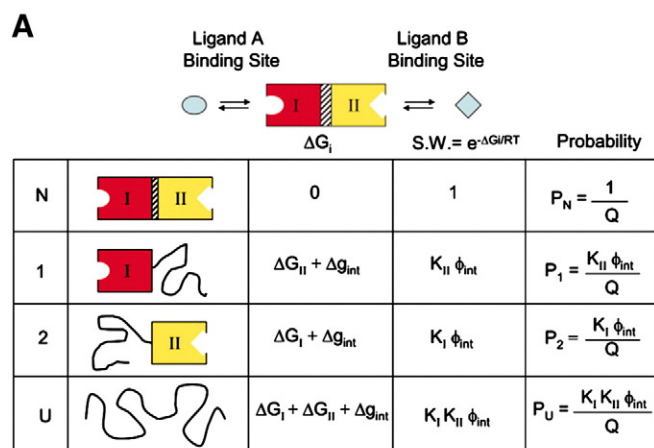
isolated to protonation at only four unique sites. To the contrary, the cooperativity arose from partial protonation effects over more than twelve positions in the protein, a result that was corroborated through experimental alanine substitution of each acidic residue in the protein (Fig. 4C).

The importance of these results for SNase is that they force reconsideration of the mechanism of pH-induced (or any ligand induced) conformational transitions, such as the Bohr-effect in hemoglobin. In the case of SNase, the effect of protonation was to perturb the stability of locally disordered states, which modulated the stability difference between the unfolded state and the ensemble of native conformations. The effect of protons could not be reconciled in the context of four unique protons binding to the unfolded conformation. Although the interpretation may not be universally applicable to all systems, this unexpected finding raises general questions regarding the allosteric mechanisms in other proteins. For example, are the observed Bohr effects in hemoglobin (or indeed *any* protein with pH-dependent stability or function) a manifestation of similar effects to the ensemble, or are the effects localized to a small

number of residues as is commonly depicted [1,62]? The analysis of SNase not only suggests that the effects may not be as originally described, but that an experimental strategy (i.e. Ala scanning at charged residues) is available that can clarify the contributions of each residue.

4. Allostery and intrinsic disorder

Allostery has emerged as especially important in the function of intrinsically disordered domains of regulatory proteins [63–65]. Consequently, the ensemble-based prediction of energy transduction in the absence of a structured pathway appears especially relevant to this class of molecules. Although a COREX/BEST analysis of the energetics of an intrinsically disordered protein is precluded by the lack of a known ground state structure, a generalization of the ensemble model has been used to derive a quantitative model for allostery in such a case [33,66]. Instead of microstates, this model treated two or three entire tethered domains within a single intrinsically disordered protein as being folded or unfolded in the presence or absence of ligand

A

	Ligand A Binding Site		ΔG_i	$S.W. = e^{-\Delta G_i/RT}$	Probability
N			0	1	$P_N = \dfrac{1}{Q}$
1			$\Delta G_{II} + \Delta g_{int}$	$K_{II}\,\phi_{int}$	$P_1 = \dfrac{K_{II}\,\phi_{int}}{Q}$
2			$\Delta G_I + \Delta g_{int}$	$K_I\,\phi_{int}$	$P_2 = \dfrac{K_I\,\phi_{int}}{Q}$
U			$\Delta G_I + \Delta G_{II} + \Delta g_{int}$	$K_I\,K_{II}\,\phi_{int}$	$P_U = \dfrac{K_I\,K_{II}\,\phi_{int}}{Q}$

B

$P_{B,Folded} \sim 0.28$

ΔG_I　ΔG_N　ΔG_2　ΔG_U

$P_{B,Folded} \sim 0.78$

ΔG_I　ΔG_N　ΔG_2　ΔG_U　ΔG_I　ΔG_N

$\Delta g_{Lig,A}$

C

Region 1

Region 2

~80-95%

~80-95%

0.15　0.10　0.07

Folding Probability of Ligand B Site (y-axis)

Folding Probability of Ligand A Site (x-axis)

(Fig. 5A). Coupling between domains was expressed as interaction energy. An exhaustive search of parameter space defined by the model objectively determined how stable the individual domains must be to maximize the allosteric communication between the domains upon binding of ligand (Fig. 5B).

Interestingly, an inverse relationship was obtained between the stability within the molecule and the allosteric coupling potential (Fig. 5C). Taken literally, this result indicated that a well-defined pathway of stable, folded structure connecting the two domains would particularly not maximize the allosteric coupling. Instead, the intrinsic disorder produced a combination of states that was optimally poised to respond to binding. Such a conclusion rationalized the experimental observation that multidomain regulatory proteins contained disproportionately high amounts of intrinsic disorder important for allosteric function [63,64].

Under the assumption that the order/disorder transitions of intrinsically disordered domains are qualitatively similar to order/disorder transitions of individual residues in a folded protein, this model developed for intrinsically disordered proteins could be extended to interpret experiments probing conformational fluctuations around the ground state structure of folded proteins (cf. Fig. 2). Such interpretation suggests that conformational fluctuations in a folded protein mediate allosteric coupling in a similar way as described above for intrinsically disordered proteins: namely that fluctuations are necessary to maximize allosteric coupling within a folded protein.

5. Functional adaptation mediated by conformational fluctuations

Recent experimental work has illuminated the importance of spontaneous fluctuations in the functional adaptation of enzyme activity. Somero and colleagues have reported that cold-adapted notothenoid fish incorporated more glycine residues in their lactate dehydrogenase sequences[67]. Their hypothesis was that the consequently increased flexibility of the native state modulated the K_m of the enzyme, allowing the cold adapted fish to survive at their physiological temperatures.

Building on this hypothesis, surface exposed single valine positions in E. coli adenylate kinase were targeted for mutation to glycine (Fig. 6A) [30,68]. These experiments tested the hypothesis that increased conformational fluctuations at these surface exposed positions would not alter the native state structure of adenylate kinase, but rather would alter the populations of excited states. The mutations were also selected with the additional criterion of being distant from residues involved in binding ligand. In effect, the mutations were designed to increase temperature-dependent local flexibility in the enzyme while preserving the integrity of both the active site and the native state structure.

The measured thermodynamics and binding affinity of the mutants revealed that the effects of these mutations on adenylate kinase were completely consistent with the original hypothesis of Somero and colleagues [30]. Namely, all of these point mutations could reduce the surrogate homeostatic temperature considerably ~8–10 °C (Fig. 6E).

Fig. 5. Energetic domain linkage in the context of intrinsic disorder: disorder maximizes site to site coupling. A. A simple ensemble model quantifying the equilibrium linkage between folding and binding of two intrinsically disordered protein domains that each bind ligand. B. Site-to-site coupling is evaluated through the addition of ligand A. Because ligand A stabilizes those states that bind A (blue arrows), the ensemble probabilities are redistributed. Binding of ligand at site A is coupled to the folding of the site B. The biologically reasonable values used for this example are: $\Delta G_I = -2.3$ kcal/mol, $\Delta G_{II} = -0.7$ kcal/mol, $\Delta g_{int} = -1.6$ kcal/mol, $\Delta g_{Lig,A} = -3.0$ kcal/mol. Note that this example is conceptually identical to that described in Fig. 2 above. C. Plot of the folding probability of domain II vs. folding probability of domain I. Colors indicate the relative coupling response upon addition of ligand (red strongest response, white is weakest response), as previously described [33]. Coupling is maximized if one or both of the domains are substantially disordered, as indicated. This figure and accompanying results are adapted from Hilser and Thompson [33]. Copyright (2007) National Academy of Sciences, USA.
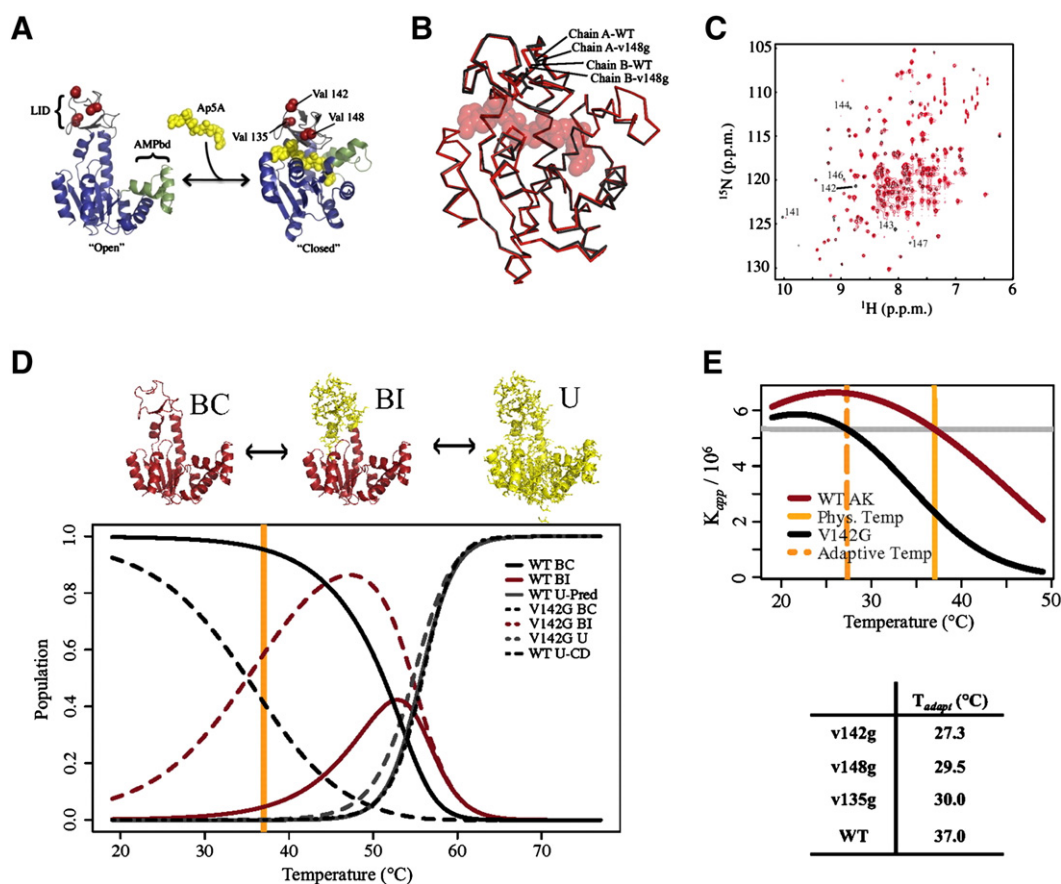
Fig. 6. Functional adaptation conferred by modulation of conformational fluctuations: rational design of surface exposed mutants in *E. coli* adenylate kinase (AK). A. Surface exposed mutations are selected to minimize possible interaction with functional sites of AK. Structures of "open" and "closed" AK are shown (PDB identifiers 4ake and 1ake, respectively) with non-hydrolyzable inhibitor Ap5A (yellow). Accepted "LID" and "AMP binding" domains are shown in gray and green, respectively. Sites of Val>Gly mutations 135, 142, and 148 are indicated in red. B. Mutant V148G does not significantly affect the crystal structure of wild-type Ap5A bound AK. All-atom RMSD values between wild-type and mutant proteins that occupy the same location in the asymmetric unit (0.2 Å) are less than the RMSD between the two copies in the same asymmetric unit (0.4 Å); PDB identifier codes 3HPQ and 3HPR. C. Mutant V142G does not significantly affect native state ^{1}H-^{15}N HSQC chemical shifts. The few residues with shifts different than wild-type are indicated with black labels. D. Surface-exposed Val>Gly mutations do modulate the populations of a new functionally relevant binding incompetent (BI) state. BI contains approximately 60 residues in the LID domain (yellow) that have undergone a local disorder transition. Estimated populations of BI are greatly increased due to Val>Gly mutations at *E. coli* physiological temperature (orange line). E. Plot of the apparent binding affinity for Ap5A calculated from isothermal titration calorimetry fitting. The K_{app} exhibited by the wild-type enzyme at its adaptive temperature of 37.0 °C is identical to the K_{app} exhibited by V142G at 27.3 °C. Point mutations can therefore decrease the functionally adapted temperature of AK by 10 °C (difference between solid and dashed orange lines). This figure and accompanying results are adapted from Schrank et al.[30].

This adaptive change in functional temperature was achieved without measurable alteration of the ground state structure of the enzyme, as verified by crystallography and NMR (Fig. 6B–C).

What was changed by the mutations was the population of excited states of the enzyme. Specifically, a locally unfolded subdomain region of approximately 60 residues encompassing the adenylate kinase "LID" domain appears to have been selectively populated by the introduction of the Gly. A new model for the transition between the open and closed states of the enzyme was thus proposed, involving selective population of the locally unfolded subdomain (Fig. 6D).

This study therefore revealed that a structure-preserving mutation approach selectively increased the probability of a locally unfolded state of adenylate kinase without changing the properties of the native fold. Since the native state structure was unchanged, the results also demonstrated that energetic perturbations (such as mutations or binding of allosteric ligands) can propagate to functional sites in the absence of a physically contacting pathway of structural distortion linking the sites. This naturally occurring design strategy appears to have decoupled the effects on function of ground state structure and spontaneous fluctuations around that structure, and represents a new paradigm for understanding mutational effects and coupling within a protein. In the context of this model, it can be seen that mutations that stabilize or destabilize a state, regardless of whether there are structural changes to the core of the protein, can affect binding or function. But rather than seeing these mutations as propagating to the active site through a pathway, such a scenario is more accurately viewed as changing the probability weighted average time the molecule spends in each state.

6. Functional residues are located in cooperatively sensitive regions of the protein

Since spontaneous fluctuations play an important role in modulating function, functional sites may also be directly coupled to fluctuations. If true, functional sites in proteins may be identifiable by how energetic perturbations at those sites affect the conformational manifold.

Liu et al. analyzed the ensemble properties of a representative set of diverse proteins, looking for such an effect (Fig. 7) [59]. The analysis was based on a metric termed "global cooperative response" (GCR) that quantified the energetic effect of a point mutation on the cooperativity between all other residue pairs in the protein. Mutations at positions that strongly affected the cooperativity of the entire protein thus exhibited high values of the GCR, and mutations that
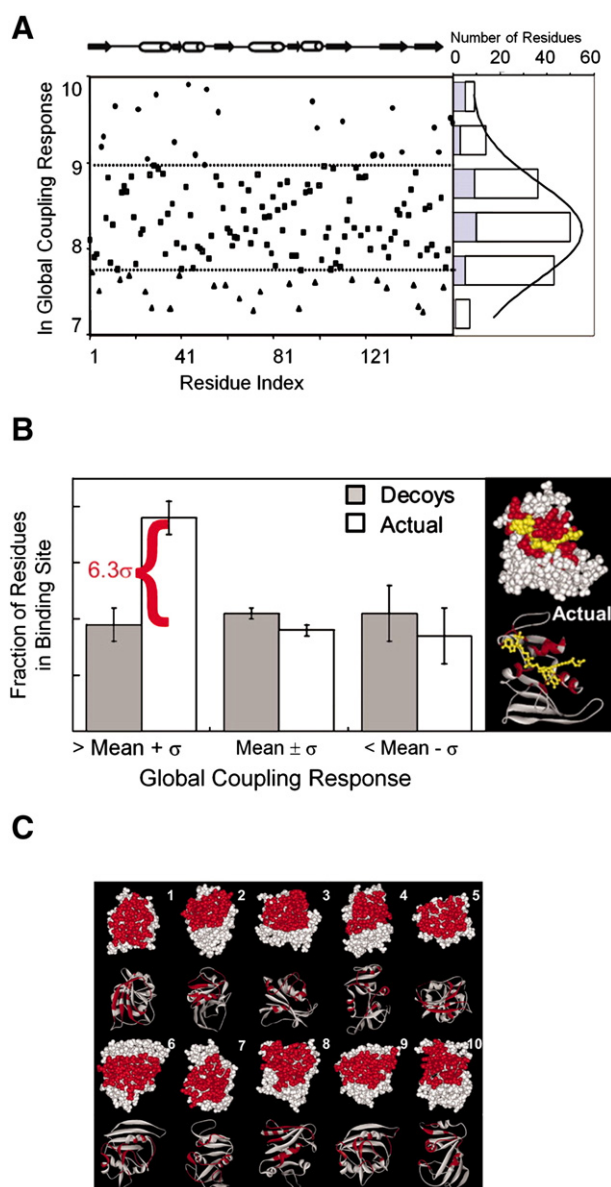
Fig. 7. Energetic hyperconnectivity of functional sites revealed by the native state ensemble. A. Global coupling response (GCR) was introduced as a measure of how an energetic perturbation (i.e. point mutation) at one residue position affects the thermodynamic linkage between all residues in a protein ensemble [59]. High GCR values suggest that the energetics of the protein ensemble are sensitive to mutation at that residue position. In *E. coli* DHFR, regions of the protein exhibiting high GCR are enriched in functional residues (blue histogram bars at the highest GCR). A substantial fraction of DHFR binding-site residues have GCR values greater than one standard deviation above the mean value for the protein. B. This enrichment is statistically significant (>6σ) as compared to GCR values observed in randomly generated DHFR binding site decoys. C. Ten examples of randomly generated binding site decoys; randomly generated sites are colored red. This figure and accompanying results are adapted from Liu et al.[59]. Copyright (2007) National Academy of Sciences, USA.

weakly affected the cooperativity exhibited low values of GCR (Fig. 7A).

A statistically significant difference in the response of the ensemble to energetic perturbations at active sites was identified (Fig. 7B–C). This indeed suggested that binding sites have been preferentially optimized to employ conformational fluctuations in carrying out function. Thus, in addition to the known importance of chemical and structural complementarity between protein and ligand, active sites appear to also exhibit a high intrinsic ability to affect the energetic coupling throughout a protein.

The analysis of binding sites raises an important question with regard to the evolution of binding sites in proteins. One possibility is that each unique protein fold is hard-wired with one or more potential active sites (i.e. a site that is hyperconnected to the rest of the protein), and thus can induce changes readily in different parts of the molecule. In this view, proteins are not merely generic scaffolds upon which arbitrary functions may be placed at arbitrary locations. Such a protein fold could be regarded as "function-competent." If this is the case, then it could indeed be possible that only those folds that are "function-competent" appear in the genomes of organisms. Those folds that are not function-competent may have been sampled, but because they could acquire no function, were purged from the genomes. An alternative possibility is that each protein fold is a "blank slate" and active sites are able to evolve in any location through mutational events. Although, the analysis presented by Liu et al. cannot unequivocally establish the actual model at play, two key pieces of information suggest that the former model of function-competent folds may be the actual evolutionary mechanism underlying the known repertoire of functional protein folds. First, the number of observed folds may be far less than the number of possible folds, based on geometric considerations [69–71]. Second, in cases where similar protein folds are found in different functional classes of proteins (which bind to different ligands), the active site is found to be in a topologically identical site [72–76], suggesting that some locations within proteins are more capable of evolving properties consistent with an active site than other locations.

7. Evolutionary conservation of spontaneous fluctuations

The importance of spontaneous fluctuations in function and adaptation has been theoretically predicted and experimentally demonstrated in model systems, demonstrating its biological importance. Such importance suggests that spontaneous fluctuations should be preserved during evolution, both at the molecular and organismal level; work from this laboratory and others [77,78] suggests that this is indeed the case.

A large scale survey of protein fold space revealed that homologous proteins, regardless of functional class or secondary structure type, generally exhibited statistically significant similarity in their residue stability constants and other thermodynamic descriptors as compared to non-homologous proteins (Fig. 8) [39,40]. Interesting exceptions to this trend were noted that suggested a thermodynamic mechanism for localized evolutionary fold change—regions of homologous proteins could change structure from alpha helix to beta sheet and still preserve the ancestral energetic characteristics[39,79].

Despite the complexity of the COREX/BEST model of the ensemble (Fig. 1), it has unexpectedly been discovered that the residue level energetic information can be largely reconstructed from sequence alone (Fig. 8) [41]. This finding has permitted development of new tools in our laboratory, such as eScape, that can characterize globular protein energetics on a genomic scale [42]. eScape uses a linear regression model, trained on the observed energetic properties of tripeptides (Fig. 9A), to estimate free energy, enthalpy, and entropy for each residue along an amino acid sequence at approximately 80% accuracy (Fig. 9B).

As there are no direct experimental methods to study the energetic properties of entire proteomes, eScape provides a unique opportunity to examine regional stability of proteomes and the energetic response of entire proteomes to selective pressure. Such a study was performed to investigate the energetic differences between the proteomes of thermophiles, mesophiles, and psychrophiles [42]. As expected, a substantial correlation between the optimal growth temperatures of the organisms and the average stabilities of their proteomes was observed (Fig. 9C). This result also implied that, in general, local fluctuations within individual proteins were relatively preserved across organisms. Interestingly, however, the extent to which individual protein families were stabilized was quite different within each
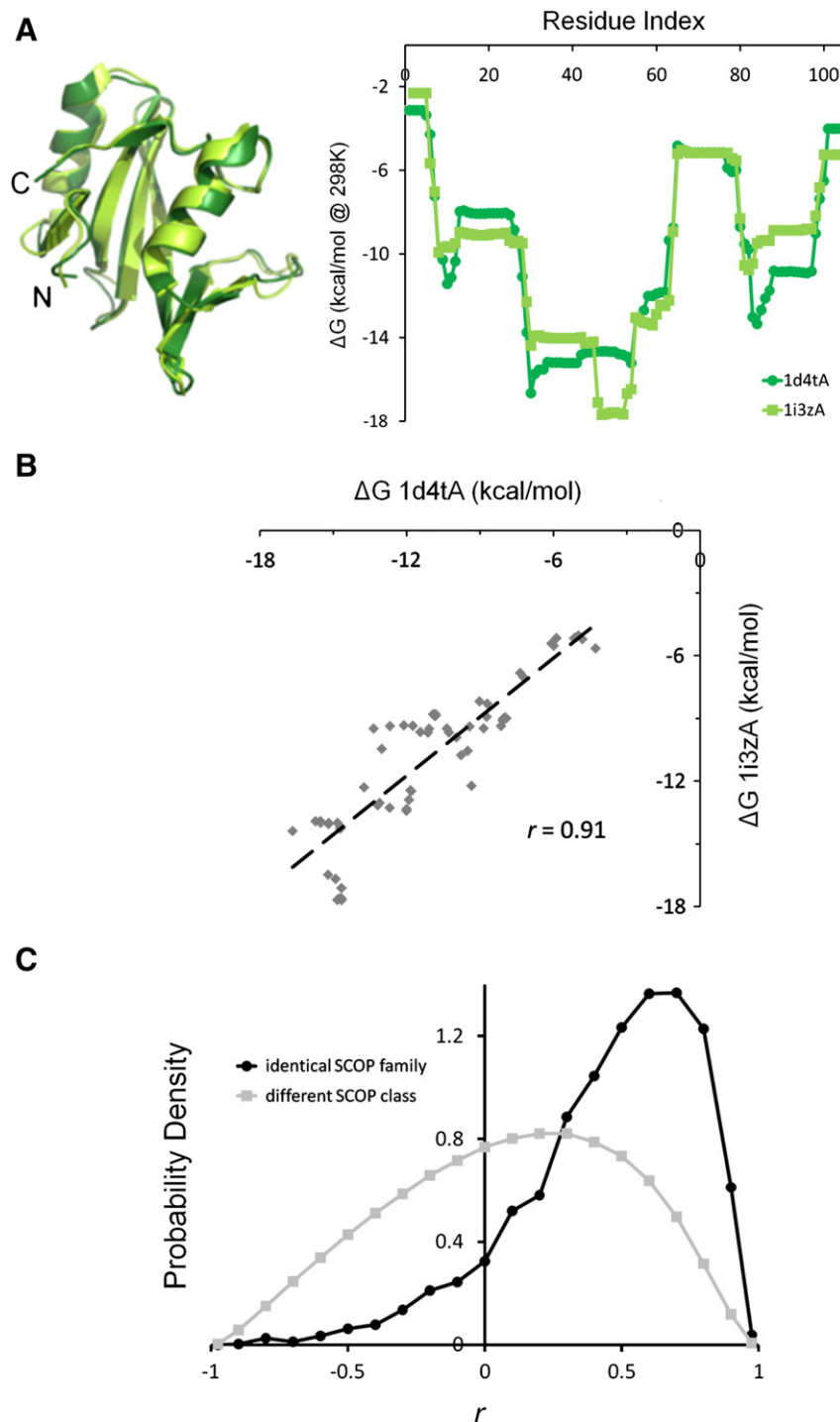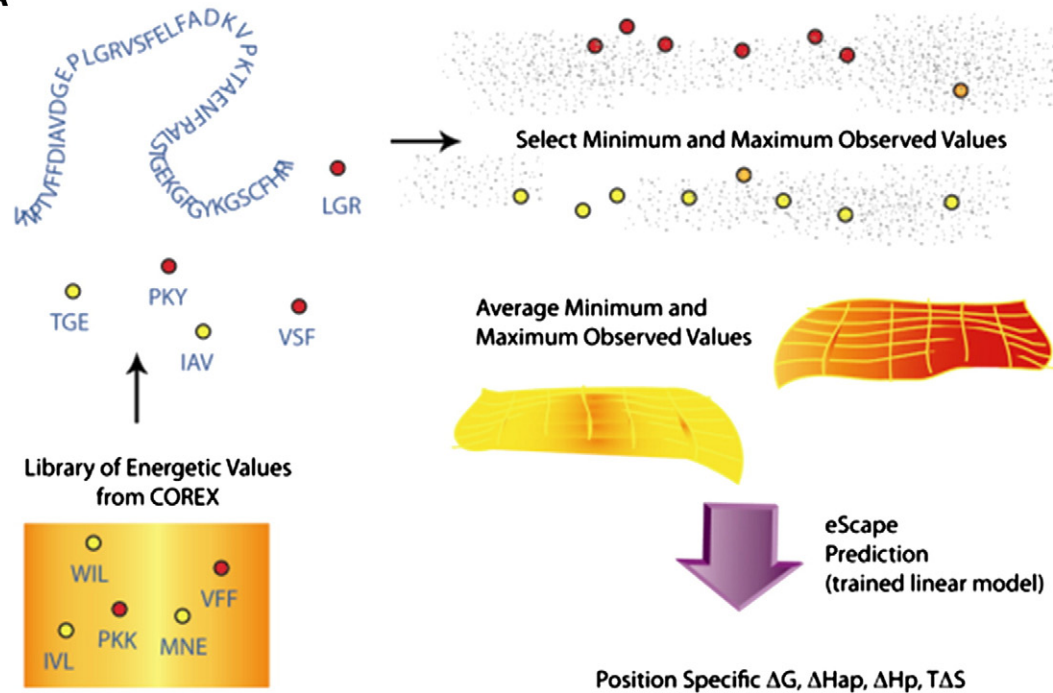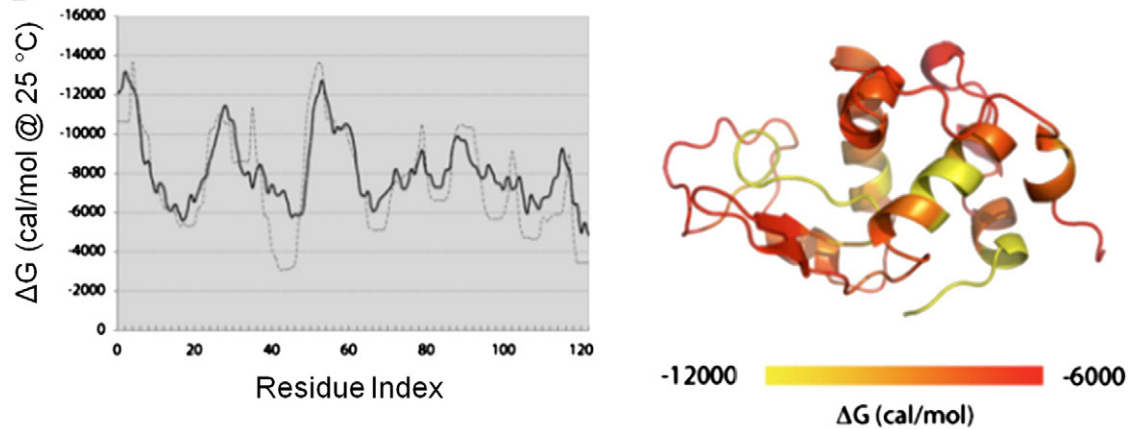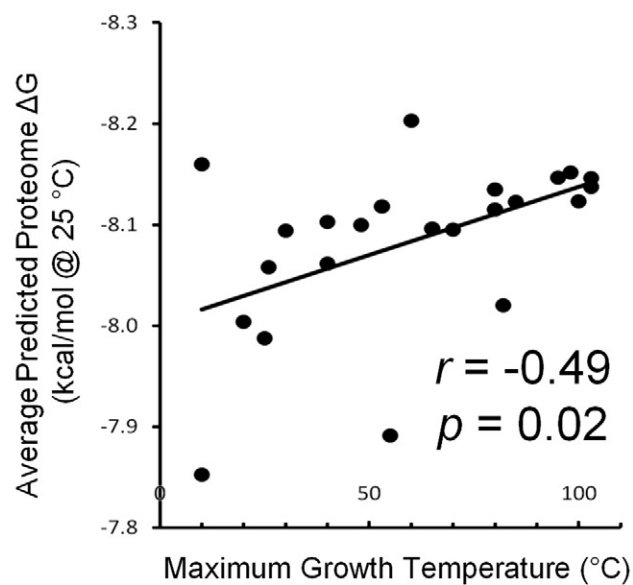
Fig. 8. Evolutionary conservation of equilibrium conformational fluctuations. A. Example of residue-specific COREX/BEST stabilities of two homologous SH2 domains, structurally aligned according to DALI [91]. B. These aligned stabilities exhibit a high Pearson correlation coefficient. C. Substantial correlations persist across a large database of pairs of homologous proteins (black curve) but are reduced when pairs of non-homologous proteins are compared (gray curve). Homology was defined according to the SCOP database [92], proteins in the same SCOP family (dark distribution) were considered homologous and proteins in different SCOP secondary structure classes (light distribution) were considered non-homologous. This figure and accompanying results are adapted from Wrabl and Hilser [39].

proteome, with enzymes and enzyme regulators being affected the most, and transcription and translation regulators being affected least [42].

8. Conclusions

Several recurring themes have emerged from more than a decade's worth of this laboratory's computational and experimental investi-

gations into the conformational fluctuations of proteins. First, diverse experimental observables (ranging from hydrogen exchange protection factors to prevalence of transcription factor intrinsic disorder to organism growth temperature) can be rationalized, unified, and predicted in many cases using a single theoretical framework based on a simple model of the protein ensemble. Second, allostery within a single protein domain, as well as signaling between multiple domains of a larger protein complex, can be explained by energetic linkage

A

Select Minimum and Maximum Observed Values

Library of Energetic Values
from COREX

Average Minimum and
Maximum Observed Values

eScape
Prediction
(trained linear model)

Position Specific ΔG, ΔHap, ΔHp, TΔS

B

C

$r = -0.49$
$p = 0.02$

without the need for a physical pathway of perturbed atoms connecting distant regions of the molecule. Third, the large degrees of conformational freedom available to a protein ensemble demands that more than one crystal structure or ground state is necessary to quantitatively understand the energetics of a protein.

The apparent conformational complexity of the protein ensemble would seem to stand in stark contrast to the visual simplicity of a single crystal structure. At a visceral level, it is somehow unappealing to accept the idea that proteins do not necessarily work as mechanical engines, complete with atomic-scaled fulcrums, ratchets, and lever arms (although there are many clear examples where they do [80–82]). However, the twin observations that conformational fluctuations are largely conserved during evolution, and that binding sites are especially cooperative with respect to the rest of the protein, strongly suggest that the ensemble is biologically essential for protein function, a notion that is shared widely among the scientific community [10,21,23,25,57, 58,83].

It is important to point out that while emphasizing the importance of dynamics and highlighting the perils of a strictly structural interpretation of proteins, it is not the intent of the work (or this review) to make the case that allostery does not often manifest itself as structural changes (which may, in some instances, even appear as a "pathway" of structural distortion). To the contrary, allosteric changes that are accompanied by structural changes are in abundance in the literature [84–87], especially in cases where there are large energetic changes, such as the light-induced isomerization of bacteriorhodopsin [88,89], or the contractile force generation of myosin [90]. The issue in question is whether an understanding of the allosteric mechanism can avail itself through an analysis of the pathways through the structure. The fact that allostery can be observed in the absence of a pathway, calls into question any quantitative accounting of allostery in terms of a pathway. Instead, the ideas emerging from this work suggest that there are organizing principles that provide a mechanistic basis by which to understand function in terms of the elementary parts of the proteins.

In the case of allostery, this is highlighted in Fig. 5. As the figure indicates, the ability to propagate the effects of the binding of ligand at one site to another site is determined by the probabilities of the different states in the ensemble. By simply changing the probabilities of states, the degree and even the sign of the allosteric coupling can be altered. The probabilities of the states, however, are a manifestation of the intrinsic stabilities (i.e. the energy difference between high and low affinity states) of the cooperative structural elements containing each binding site and the coupling energy between the two sites. We propose that allostery can be understood in the context of this model. In other words, experimental access to the conformational free energies and interaction energies can provide not just a qualitative understanding of the mechanism, but a quantitative understanding of the impact of mutations. Importantly, because the model accounts for allostery in terms of the local stabilities and interaction terms, the effects of surface-exposed residues and residues distal to either binding site can, in principle, be equally well described by this approach. In this respect, the current ensemble allosteric model provides a means of identifying similar thermodynamic architecture (e.g. where the stabilities and interaction energies are in the same region of thermodynamic parameter space in Fig. 5) between different proteins.

A future challenge is practical use of the protein ensemble in design of structure or function. Although this seems a daunting problem, clues for its solution come from both natural and laboratory examples of the effects of distal mutations on binding affinities in DHFR and AK. Indeed, rationally designed surface exposed mutations in AK (a strategy already documented in nature) appear to decouple effects on excited states from effects on the ground state. Preliminary unpublished data from this laboratory point to the exciting possibility that such mutations result in nuanced effects that allow us to rationally decouple K_m from k_{cat}. These first steps encourage optimism that protein design, via a strategy that "sculpts the ensemble" to enhance enzymatic function, binding affinity, or thermal stability, could someday be a reality.

Acknowledgments

References

[1] A.L. Lehninger, Biochemistry, 2nd ed. Worth, New York, 1975.

[2] A.R. Fersht, Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding, W.H. Freeman and Company, New York, 1999.

[3] T.L. Creighton, Proteins: Structures and Molecular Properties, 2 ed. W.H. Freeman and Company, New York, 1993.

[4] K. Henzler-Wildman, D. Kern, Dynamic personalities of proteins, Nature 450 (2007) 964–972.

[5] N. Tokuriki, D.S. Tawfik, Protein dynamism and evolvability, Science (New York, N.Y) 324 (2009) 203–207.

[6] P.N. Bryan, J. Orban, Proteins that switch folds, Current Opinion in Structural Biology, 2010.

[7] S.J. Benkovic, S. Hammes-Schiffer, A perspective on enzyme catalysis, Science (New York, N.Y) 301 (2003) 1196–1202.

[8] Y. Bai, T.R. Sosnick, L. Mayne, S.W. Englander, Protein folding intermediates: native-state hydrogen exchange, Science (New York, N.Y) 269 (1995) 192–197.

[9] A. Mittermaier, L.E. Kay, Observing biological dynamics at atomic resolution using NMR, Trends in Biochemical Sciences 34 (2009) 601–611.

[10] D. Kern, E.R. Zuiderweg, The role of dynamics in allosteric regulation, Current Opinion in Structural Biology 13 (2003) 748–757.

[11] K.A. Henzler-Wildman, M. Lei, V. Thai, S.J. Kerns, M. Karplus, D. Kern, A hierarchy of timescales in protein dynamics is linked to enzyme catalysis, Nature 450 (2007) 913–916.

[12] F. Stengel, A.J. Baldwin, A.J. Painter, N. Jaya, E. Basha, L.E. Kay, E. Vierling, C.V. Robinson, J.L. Benesch, Quaternary dynamics and plasticity underlie small heat shock protein chaperone function, Proceedings of the National Academy of Sciences of the United States of America 107 (2010) 2007–2012.

[13] V.A. Feher, J. Cavanagh, Millisecond-timescale motions contribute to the function of the bacterial response regulator protein Spo0F, Nature 400 (1999) 289–293.

[14] K.K. Frederick, M.S. Marlow, K.G. Valentine, A.J. Wand, Conformational entropy in molecular recognition by proteins, Nature 448 (2007) 325–328.

[15] P.J. Sapienza, A.L. Lee, Using NMR to study fast dynamics in proteins: methods and applications, Current Opinion in Pharmacology 10 (2010) 723–730.

[16] C.M. Petit, J. Zhang, P.J. Sapienza, E.J. Fuentes, A.L. Lee, Hidden dynamic allostery in a PDZ domain, Proceedings of the National Academy of Sciences of the United States of America 106 (2009) 18249–18254.

[17] J. Monod, J. Wyman, J.P. Changeaux, On the nature of allosteric transitions: a plausible model, Journal of Molecular Biology 12 (1965) 88–118.

[18] D.E. Koshland, G. Nemethy, D. Filmer, Comparison of experimental binding data and theoretical models in proteins containing subunits, Biochemistry 5 (1966) 365–385.

[19] J.A. Boyer, C.J. Clay, K.S. Luce, M.H. Edgell, A.L. Lee, Detection of native state nonadditivity in double mutant cycles via hydrogen exchange, Journal of the American Chemical Society 132 (2010) 8010–8019.

Fig. 9. The eScape algorithm estimates protein energetics from sequence information and is applicable on a proteomic scale. A. The eScape algorithm is parameterized from tripeptide information contained in a large database of proteins analyzed with the COREX/BEST algorithm. B. The eScape algorithm recapitulates the observed residue-specific stabilities of human α-lactalbumin, as well as most other proteins tested, at 80% accuracy. The x-axis of the plot represents the residue position index of human α-lactalbumin (PDB 1b9o) and the y-axis represents the predicted free energy of stability of the folded structure as reported at that residue position. The dark line denotes eScape results and the lighter line the COREX/BEST results for the same protein. Notably, the agreement between the two lines implies that equilibrium conformational fluctuations of proteins are largely encoded by local sequence information. C. A significant Pearson correlation coefficient is observed between the average eScape stabilities of entire proteomes and the maximum growth temperatures of 23 diverse microbes. Each point in the scatterplot represents data from one microbe. Data, figures, and accompanying results are adapted from Gu and Hilser [41,42].

[20] J.C. Lee, M. Natarajan, V. Nashine, M. Socolich, T. Vu, W.P. Russ, S.J. Benkovic, R. Ranganathan, Surface sites for engineering allosteric control in proteins, Science (New York, N.Y) 322 (2008) 438–442.

[21] C.J. Tsai, A. del Sol, R. Nussinov, Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms, Molecular BioSystems 5 (2009) 207–216.

[22] R. Elber, Simulations of allosteric transitions, Current Opinion in Structural Biology 21 (2011) 167–172.

[23] S.R. Tzeng, C.G. Kalodimos, Protein dynamics and allostery: an NMR view, Current Opinion in Structural Biology 21 (2011) 62–67.

[24] P. Csermely, R. Palotai, R. Nussinov, Induced fit, conformational selection and independent dynamic segments: an extended view of binding events, Trends in Biochemical Sciences 35 (2010) 539–546.

[25] K. Gunasekaran, B. Ma, R. Nussinov, Is allostery an intrinsic property of all dynamic proteins? Proteins: Struct Funct Bioinf 57 (2004) 433–443.

[26] V.J. Hilser, E. Freire, Structure-based calculation of the equilibrium folding pathway of proteins. Correlation with hydrogen exchange protection factors, Journal of Molecular Biology 262 (1996) 756–772.

[27] V.J. Hilser, J. Gomez, E. Freire, The enthalpy change in protein folding and binding: refinement of parameters for structure-based calculations, Proteins 26 (1996) 123–133.

[28] J. Gomez, V.J. Hilser, D. Xie, E. Freire, The heat capacity of proteins, Proteins 22 (1995) 404–412.

[29] V.J. Hilser, Modeling the native state ensemble, Methods in Molecular Biology (Clifton, N.J) 168 (2001) 93–116.

[30] T.P. Schrank, D.W. Bolen, V.J. Hilser, Rational modulation of conformational fluctuations in adenylate kinase reveals a local unfolding mechanism for allostery and functional adaptation in proteins, Proceedings of the National Academy of Sciences of the United States of America 106 (2009) 16984–16989.

[31] S.T. Whitten, B.E. Garcia-Moreno, V.J. Hilser, Ligand effects on the protein ensemble: unifying the descriptions of ligand binding, local conformational fluctuations, and protein stability, Methods in Cell Biology 84 (2008) 871–891.

[32] K. Sayar, O. Ugur, T. Liu, V.J. Hilser, O. Onaran, Exploring allosteric coupling in the alpha-subunit of heterotrimeric G proteins using evolutionary and ensemble-based approaches, BMC Structural Biology 8 (2008) 23.

[33] V.J. Hilser, E.B. Thompson, Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins, Proceedings of the National Academy of Sciences of the United States of America 104 (2007) 8311–8315.

[34] S.T. Whitten, A.J. Kurtz, M.S. Pometun, A.J. Wand, V.J. Hilser, Revealing the nature of the native state ensemble through cold denaturation, Biochemistry 45 (2006) 10163–10174.

[35] T. Liu, S.T. Whitten, V.J. Hilser, Ensemble-based signatures of energy propagation in proteins: a new view of an old phenomenon, Proteins 62 (2006) 728–738.

[36] V.J. Hilser, E.B. Garcia-Moreno, T.G. Oas, G. Kapp, S.T. Whitten, A statistical thermodynamic model of the protein ensemble, Chemical Reviews 106 (2006) 1545–1558.

[37] H. Pan, J.C. Lee, V.J. Hilser, Binding sites in Escherichia coli dihydrofolate reductase communicate by modulating the conformational ensemble, Proceedings of the National Academy of Sciences of the United States of America 97 (2000) 12020–12025.

[38] V.J. Hilser, D. Dowdy, T.G. Oas, E. Freire, The structural distribution of cooperative interactions in proteins: analysis of the native state ensemble, Proceedings of the National Academy of Sciences of the United States of America 95 (1998) 9903–9908.

[39] J.O. Wrabl, V.J. Hilser, Investigating homology between proteins using energetic profiles, PLoS Computational Biology 6 (2010) e1000722.

[40] J. Vertrees, J.O. Wrabl, V.J. Hilser, Energetic profiling of protein folds, Methods in Enzymology 455 (2009) 299–327.

[41] J. Gu, V.J. Hilser, Predicting the energetics of conformational fluctuations in proteins from sequence: a strategy for profiling the proteome, Structure 16 (2008) 1627–1637.

[42] J. Gu, V.J. Hilser, Sequence-based analysis of protein energy landscapes reveals nonuniform thermal adaptation within the proteome, Molecular Biology and Evolution 26 (2009) 2217–2227.

[43] J. Vertrees, P. Barritt, S. Whitten, V.J. Hilser, COREX/BEST server: a web browser-based program that calculates regional stability variations within protein structures, Bioinformatics (Oxford, England) 21 (2005) 3318–3319.

[44] A.J. Doig, Recent advances in helix-coil theory, Biophysical Chemistry 101–102 (2002) 281–293.

[45] K.P. Murphy, E. Freire, Thermodynamics of structural stability and cooperative folding behavior in proteins, Advances in Protein Chemistry 43 (1992) 313–361.

[46] K.P. Murphy, S.J. Gill, Solid model compounds and the thermodynamics of protein unfolding, Journal of Molecular Biology 222 (1991) 699–709.

[47] K.P. Murphy, P.L. Privalov, S.J. Gill, Common features of protein unfolding and dissolution of hydrophobic compounds, Science (New York, N.Y) 247 (1990) 559–561.

[48] J.A. D'Aquino, J. Gomez, V.J. Hilser, K.H. Lee, L.M. Amzel, E. Freire, The magnitude of the backbone conformational entropy change in protein folding, Proteins 25 (1996) 143–156.

[49] D.A. McQuarrie, J.D. Simon, Molecular Thermodynamics, University Science Books, Sausalito, CA, 1999.

[50] J.O. Wrabl, S.A. Larson, V.J. Hilser, Thermodynamic environments in proteins: fundamental determinants of fold specificity, Protein Science 11 (2002) 1945–1957.

[51] J.O. Wrabl, S.A. Larson, V.J. Hilser, Thermodynamic propensities of amino acids in the native state ensemble: implications for fold recognition, Protein Science 10 (2001) 1032–1045.

[52] J.O. Wooll, J.O. Wrabl, V.J. Hilser, Ensemble modulation as an origin of denaturant-independent hydrogen exchange in proteins, Journal of Molecular Biology 301 (2000) 247–256.

[53] C.R. Babu, V.J. Hilser, A.J. Wand, Direct access to the cooperative substructure of proteins and the protein ensemble via cold denaturation, Nature Structural & Molecular Biology 11 (2004) 352–357.

[54] S.T. Whitten, E.B. Garcia-Moreno, V.J. Hilser, Local conformational fluctuations can modulate the coupling between proton binding and global structural transitions in proteins, Proceedings of the National Academy of Sciences of the United States of America 102 (2005) 4282–4287.

[55] A. Manson, S.T. Whitten, J.C. Ferreon, R.O. Fox, V.J. Hilser, Characterizing the role of ensemble modulation in mutation-induced changes in binding affinity, Journal of the American Chemical Society 131 (2009) 6785–6793.

[56] S.A. Larson, V.J. Hilser, Analysis of the "thermodynamic information content" of a Homo sapiens structural database reveals hierarchical thermodynamic organization, Protein Science 13 (2004) 1787–1801.

[57] A. Cooper, D.T.F. Dryden, Allostery without conformational change, European Biophysics Journal 11 (1984) 105–109.

[58] Q. Cui, M. Karplus, Allostery and cooperativity revisited, Protein Science 17 (2008) 1295–1307.

[59] T. Liu, S.T. Whitten, V.J. Hilser, Functional residues serve a dominant role in mediating the cooperativity of the protein ensemble, Proceedings of the National Academy of Sciences of the United States of America 104 (2007) 4347–4352.

[60] M.J. Osborne, J. Schnell, S.J. Benkovic, H.J. Dyson, P.E. Wright, Backbone dynamics in dihydrofolate reductase complexes: role of loop flexibility in the catalytic mechanism, Biochemistry 40 (2001) 9846–9859.

[61] C.A. Fierke, K.A. Johnson, S.J. Benkovic, Construction and evaluation of the kinetic scheme associated with dihydrofolate reductase from Escherichia coli, Biochemistry 26 (1987) 4085–4092.

[62] J.M. Berg, J.L. Tymoczko, L. Stryer, Biochemistry, 6th ed. W.H. Freeman and Company, New York, 2007.

[63] J. Liu, N.B. Perumal, C.J. Oldfield, E.W. Su, V.N. Uversky, A.K. Dunker, Intrinsic disorder in transcription factors, Biochemistry 45 (2006) 6873–6888.

[64] Y. Minezaki, K. Homma, A.R. Kinjo, K. Nishikawa, Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation, Journal of Molecular Biology 359 (2006) 1137–1149.

[65] A.S. Garza, N. Ahmad, R. Kumar, Role of intrinsically disordered protein regions/domains in transcriptional regulation, Life Sciences 84 (2009) 189–193.

[66] I. Luque, S.A. Leavitt, E. Freire, The linkage between protein folding and functional cooperativity: two sides of the same coin? Annual Review of Biophysics and Biomolecular Structure 31 (2002) 235–256.

[67] P.A. Fields, G.N. Somero, Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenoid fishes, Proceedings of the National Academy of Sciences of the United States of America 95 (1998) 11476–11481.

[68] T.P. Schrank, W.A. Elam, J. Li, V.J. Hilser, Strategies for the thermodynamic characterization of linked binding/local folding reactions within the native state: application to the LID domain of adenylate kinase from Escherichia coli, Methods in Enzymology 492 (2011) 253–282.

[69] B. Grainger, M.I. Sadowski, W.R. Taylor, Re-evaluating the "rules" of protein topology, Journal of Computational Biology 17 (2010) 1371–1384.

[70] H. Li, C. Tang, N.S. Wingreen, Are protein folds atypical? Proceedings of the National Academy of Sciences of the United States of America 95 (1998) 4987–4990.

[71] P. Cossio, A. Trovato, F. Pietrucci, F. Seno, A. Maritan, A. Laio, Exploring the universe of protein structures beyond the protein data bank, PLoS Computational Biology 6 (2010) e1000957.

[72] N. Nagano, C.A. Orengo, J.M. Thornton, One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions, Journal of Molecular Biology 321 (2002) 741–765.

[73] A.N. Lupas, C.P. Ponting, R.B. Russell, On the evolution of protein folds: are similar motifs in different protein folds the result of convergence, insertion, or relics of an ancient peptide world? Journal of Structural Biology 134 (2001) 191–203.

[74] A.G. Murzin, S.E. Brenner, T. Hubbard, C. Chothia, SCOP: a structural classification of proteins database for the investigation of sequences and structures, Journal of Molecular Biology 247 (1995) 536–540.

[75] Y. Qi, N.V. Grishin, Structural classification of thioredoxin-like fold proteins, Proteins: Structure, Function, and Bioinformatics 58 (2005) 376–388.

[76] R.B. Russell, P.D. Sasieni, M.J.E. Sternberg, Supersites within superfolds: binding site similarity in the absence of homology, Journal of Molecular Biology 282 (1998) 903–918.

[77] A.B. Law, E.J. Fuentes, A.L. Lee, Conservation of side-chain dynamics within a protein family, Journal of the American Chemical Society 131 (2009) 6322–6333.

[78] S.M. Hollup, E. Fuglebakk, W.R. Taylor, N. Reuter, Exploring the factors determining the dynamics of different protein folds, Protein Science 20 (2011) 197–209.

[79] J. Vertrees, J.O. Wrabl, V.J. Hilser, An energetic representation of protein architecture that is independent of primary and secondary structure, Biophysical Journal 97 (2009) 1461–1470.

[80] J.A. Spudich, S. Sivaramakrishnan, Myosin VI: an innovative motor that challenged the swinging lever arm hypothesis, Nature Reviews Molecular Cell Biology 11 (2010) 128–137.

[81] A.L. Horwich, G.W. Farr, W.A. Fenton, GroEL-GroES-mediated protein folding, Chemical Reviews 106 (2006) 1917–1930.

[82] T.A. Steitz, The structural changes of T7 RNA polymerase from transcription initiation to elongation, Current Opinion in Structural Biology 19 (2009) 683–690.

[83] M.J. Whitley, A.L. Lee, Frameworks for understanding long-range intra-protein communication, Current Protein Peptide Science 10 (2009) 116–127.

[84] G.K. Ackers, Deciphering the molecular code of hemoglobin allostery, Advances in Protein Chemistry 51 (1998) 185–253.

[85] E. DiCera, Thrombin, Molecular Aspects of Medicine 29 (2008) 203–254.

[86] J.A. Hardy, J.A. Wells, Searching for new allosteric sites in enzymes, Current Opinion in Structural Biology 14 (2004) 706–715.

[87] P. Hauske, C. Ottmann, M. Meltzer, M. Ehrmann, M. Kaiser, Allosteric regulation of proteases, ChemBioChem 9 (2008) 2920–2928.

[88] U. Haupts, J. Tittor, D. Oesterhelt, Closing in on bacteriorhodopsin: progress in understanding the molecule, Annual Review of Biophysics and Biomolecular Structure 28 (1999) 367–399.

[89] F. Gai, K.C. Hasson, J.C. McDonald, P.A. Anfinrud, Chemical dynamics in proteins: the photoisomerization of retinal in bacteriorhodopsin, Science (New York, NY) 279 (1998) 1886–1891.

[90] M. Tyska, D.M. Warshaw, The myosin power stroke, Cell Motility and the Cytoskeleton 51 (2002) 1–15.

[91] L. Holm, J. Park, DaliLite workbench for protein structure comparison, Bioinformatics (Oxford, England) 16 (2000) 566–567.

[92] A. Andreeva, D. Howorth, J.M. Chandonia, S.E. Brenner, T.J. Hubbard, C. Chothia, A.G. Murzin, Data growth and its impact on the SCOP database: new developments, Nucleic Acids Research 36 (2008) D419–D425.