ACCOUNTS of chemical research

Article

pubs.acs.org/accounts

Fluctuations within Folded Proteins: Implications for Thermodynamic and Allosteric Regulation

Published as part of the Accounts of Chemical Research special issue "Protein Motion in Catalysis".

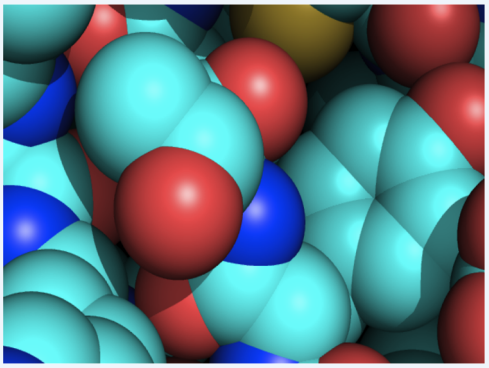
Kateri H. DuBay,[†] Gregory R. Bowman,[‡] and Phillip L. Geissler*[,§]

[†]Department of Chemistry, University of Virginia, Charlottesville, Virginia 22904, United States

[‡]Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110, United States

[§]Department of Chemistry, University of California, Berkeley, Berkeley, California 94720, United States

CONSPECTUS: Folded protein structures are both stable and dynamic. Historically, our clearest window into these structures came from X-ray crystallography, which generally provided a static image of each protein's singular "folded state", highlighting its stability. Deviations away from that crystallographic structure were difficult to quantify, and as a result, their potential functional consequences were often neglected. However, several dynamical and statistical studies now highlight the structural variability that is present within the protein's folded state. Here we review mounting evidence of the importance of these structural rearrangements; both experiment and computation indicate that folded proteins undergo substantial fluctuations that can greatly influence their function.



Crucially, recent studies have shown that structural elements of proteins, especially their side-chain degrees of freedom, fluctuate in ways that generate significant conformational heterogeneity. The entropy associated with these motions contributes to the folded structure's thermodynamic stability. In addition, since these fluctuations can shift in response to perturbations such as ligand binding, they may play an important role in the protein's capacity to respond to environmental cues. In one compelling example, the entropy associated with side-chain fluctuations contributes significantly to regulating the binding of calmodulin to a set of peptide ligands.

The neglect of fluctuations within proteins' native states was often justified by the dense packing within folded proteins, which has inspired comparisons with crystalline solids. Many liquids, however, can achieve similarly dense packing yet fluidity is maintained through correlated molecular motions. Indeed, the studies we discuss favor comparison of folded proteins not with solids but instead with dense liquids, where the internal side chain fluidity is facilitated by collective motions that are correlated over long distances. These correlated rearrangements can enable allosteric communication between different parts of a protein, through subtle and varied channels. Such long-range correlations appear to be an innate feature of proteins in general, manifest even in molecules lacking known allosteric regulators and arising robustly from the physical nature of their internal environment. Given their ubiquity, it is only to be expected that, over time, nature has refined some subset of these correlated motions and put them to use.

Native state fluctuations increasingly appear to be vital for proteins' natural functions. Understanding the diversity, origin, and range of these rearrangements may provide novel routes for rationally manipulating biomolecular activity.

1. INTRODUCTION

Living organisms, in their quest to survive and propagate within diverse and ever-changing environments, must achieve two oft-contradictory goals: stability and adaptability. Both are required for organisms to self-organize, process energy, grow, and reproduce within environments that are anything but static. These life processes unfold via complex webs of interactions between complex sets of material components, and thus the interactions and materials that compose such organisms often themselves encompass these rival abilities both to change and to remain the same.

The protein, one of the fundamental building blocks of known life, exemplifies this duality. Despite being made up of a remarkably small set of conserved components, a great variety of proteins, diverse in both structure and function, are to be found: their encryption within stable genetic material has allowed individual proteins to propagate relatively unchanged over vast stretches of time, while at the same time the diversity of the current proteome attests to their historical ability to adapt in the face of new needs and conditions.

Received: September 27, 2014
Published: February 17, 2015

DOI: 10.1021/ar500351b
Acc. Chem. Res. 2015, 48, 1098−1105

Within the context of our current proteome, natively folded and catalytically active proteins provide particularly good examples of this duality within a single biostructure. Stably folded, yet amenable to regulation, they often participate in robust interaction networks designed explicitly to alter their function in response to environmental cues. Within their own structure, these well-folded proteins embody a similar dichotomy of stability and flexibility that facilitates these functions. These proteins are densely folded (with densities similar to those of organic crystals) and provide well-resolved and reproducible crystal structures when probed. However, as we and others have recently highlighted, at physiological temperatures they also contain significant internal fluctuations.[1−5]

Fluctuations within such densely packed systems of necessity take place through correlated motions involving the concerted rearrangements of various components, for example, the movement of one side chain in a protein may facilitate the movement of another, distantly located, side chain as it fluctuates within and between the local minima that dot its conformational free energy landscape. These fluctuations and their correlated nature have been largely neglected when thinking about the catalytic activity of folded proteins. Their presence, however, suggests some intriguing possibilities, providing potential entropic and allosteric mechanisms that may assist in regulating protein function and catalysis. Although their contributions are often difficult to assess, both effects are likely to contribute to a protein's ability to reliably assist in bioreactions and to do so in an environmentally responsive manner.

In this Account, we review recent findings from our group, consider how they fit in with those of others, and discuss their broader implications for the field of protein regulation. In section 2, we first describe in more detail the significant fluctuations found even within the cores of stable, natively folded proteins (section 2.1).[5] We then focus on the fluctuations found among side chains within these systems and highlight their potential entropic contributions to the thermodynamic regulation of protein interactions (section 2.2).[4,6] In section 3, we take a look at the correlations expected to arise in such a densely packed but fluctuating system and explore the effect of those correlations in terms of allosteric regulation, wherein the functional behavior of the protein at one location is modified by some perturbation, such as the binding of a signaling molecule, at another location. We first examine the role of cryptic allostery (section 3.1)[7] and then consider the extent of side-chain correlations that are possible even in the absence of backbone motion (section 3.2).[8] Finally, in section 4, we conclude with a discussion of how the fluctuations that we have highlighted within the natively folded protein contribute to both its stability and its adaptability within the ever-changing environments that form the backdrop to biological life.

2. PROTEIN FLUCTUATIONS AND THERMODYNAMIC REGULATION

The presence of certain kinds of motion has long been recognized in those proteins where movement is readily observable and essential for function, such as the gear-like spinning of ATP synthases,[9] the walking strides of kinesins,[10] and even the concerted rearrangements involved in the binding of oxygen molecules to hemoglobin.[11] In these cases, and others like them, the observed motion involves coordinated, large-scale rearrangements of the protein's backbone and often occurs in a cyclic or repetitive manner, greatly facilitating its detection and study.

Our focus here, however, is on the well-folded proteins that have historically been thought of as fairly static. These proteins, as we discuss in more detail below, vary far more about their idealized folded structures than this traditional picture suggests, particular among their side chains. Aside from the motions of a few key residues in the binding pockets of specific proteins during catalysis, the influence of these more subtle fluctuations on protein function has been largely ignored and understandably so, as these fluctuations generally occur on smaller length and time scales, making them more challenging to observe and quantify and their biological purpose, if any, more difficult to discern. But despite their smaller scales, their potential influence on protein interactions is large. Substantial amounts of entropy have been observed in isothermal calorimetry experiments,[12−14] and at times, these entropic contributions can even outweigh an unfavorable enthalpy of binding.[15] Recent work on the catabolite activator protein has highlighted the vital role played by these entropic contributions in ligand binding.[16] In this section, we explore both the extent of these fluctuations and their potential contributions to the thermodynamic regulation of protein interactions.

2.1. Significant Motion within Folded Proteins

The idea that proteins are not rigid bodies has a long history. Early NMR work demonstrated that even large aromatic rings can flip within the hydrophobic cores of proteins.[17] Hydrogen−deuterium exchange experiments also detected more substantial fluctuations away from the native state.[18] Consistent with these experimental results, atomically detailed molecular dynamics simulations captured some of the variability of protein structures,[19−21] though complete ring flips were beyond the reach of existing sampling techniques. Theorists developed models that demonstrated that allosteric signals could be conveyed by changes in the distribution of structures accessible to regions of a protein.[22]

Later work more fully embraced the idea of protein flexibility and began tying it to the functions that proteins perform. For example, measurements of side-chain and backbone NMR order parameters revealed that side-chain conformations are more variable than backbone conformations.[1] Measuring these order parameters in the absence and presence of ligands demonstrated that ligand binding can induce entropic changes in distant portions of a protein and, therefore, contribute to allosteric communication.[14] NMR experiments on enzymes also found that the rates of protein conformational changes are correlated with catalytic rates.[23,24] While apparently at odds with the single structures solved via X-ray crystallography, recent X-ray studies at room temperature have revealed substantially more flexibility than is observed at cryogenic temperatures and demonstrated that these alternate structures are important for allosteric communication.[2] Advances in computational modeling also revealed substantially more flexibility than had previously been observed and highlighted the implications of this variability for allosteric communication and catalysis.[3,25−28]

Our own work has built on this foundation by leveraging new computational methods to dramatically increase the time scales that we can address and draw closer to experiment. While conventional molecular dynamics simulations are typically limited to capturing microsecond time scales, we capture hundreds of microseconds to millisecond time scale events by building Markov state models from thousands of independent simulations.[29,30] These models work by identifying the structural states where a protein tends to dwell, the equilibrium populations

DOI: 10.1021/ar500351b
Acc. Chem. Res. 2015, 48, 1098−1105

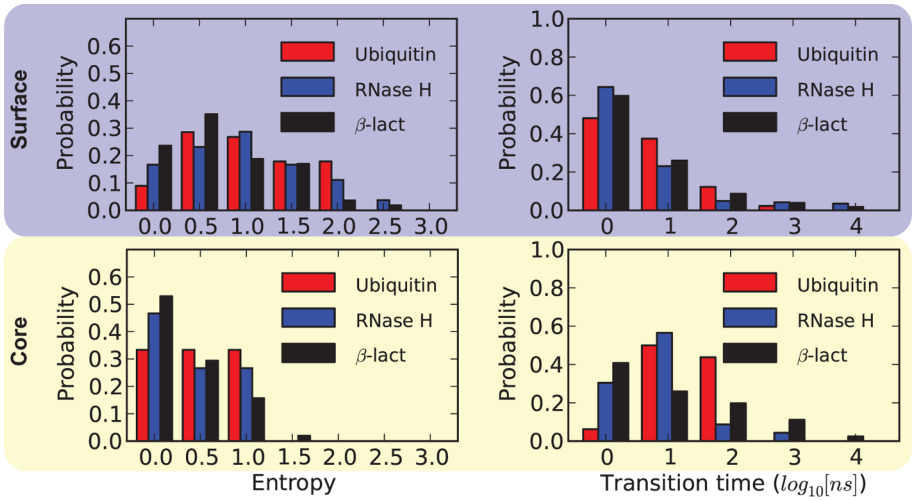


Figure 1. Histograms of the entropies of side-chain rotameric states (left) and the time scales for transitioning between rotamers (right) for surface residues (top, blue) and core residues (bottom, yellow). β-lact is β-lactamase. Reproduced with permission from ref 5. Copyright 2014 American Chemical Society.

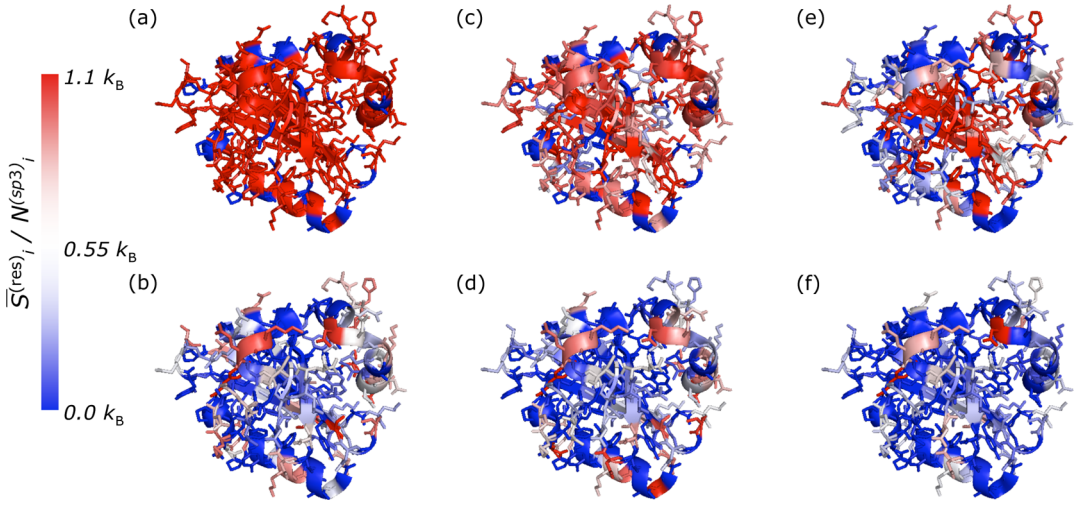


Figure 2. Side-chain conformational entropies in the photoactive yellow protein (1F9I[34]), as influenced by various interatomic forces. Entropy values were normalized by the number of sp^3-sp^3 hybridized rotatable bonds in each residue, such that red indicates those residues having a maximum possible entropy while blue indicates those with a minimum. The combinations of interatomic forces considered included (a) a noninteracting reference system (blue residues indicate amino acids without rotatable bonds), (b) Lennard-Jones (LJ) interactions, (c) implicit solvent (IS) interactions, (d) both LJ and IS interactions, (e) hydrogen-bonding (HB) and salt-bridge (SB) interactions, and (f) all of the above. Reproduced with permission from ref 4. Copyright 2009 Elsevier.

of these states, and the probabilities of transitioning between them. Using this information, we systematically examined the conformational heterogeneity (Shannon entropy) of the backbone and side chains of a set of three proteins.[5] This analysis revealed substantial heterogeneity, even within the densely packed cores of each protein (Figure 1). In particular, we found that basically every side chain visits multiple rotameric states and that simulating longer time scales reveals substantially more heterogeneity than is observed on very short time scales. These results suggest that proteins are more liquid-like than crystalline. To ensure that this heterogeneity is not erroneous, we also calculated NMR order parameters for both the backbone and side chains and demonstrated reasonable agreement with existing experimental measurements. For example, the correlation coefficients between simulation and experiment are often between 0.6 and 0.8. There are some cases where correlation coefficients are worse (e.g., 0.3), but this is typically when the experimental error bars are larger than 0.1−0.2. There are also cases where the average order parameters in simulation and experiment agree well but the correlation coefficient is small because the spread in the data is tiny and even small errors dramatically reduce the correlation. In the future, we expect improved force fields and sampling techniques to increase the agreement between simulation and experiment. Better agreement may also come from understanding the limits of experimental measurements (e.g., the time scales to which they are sensitive to) and either mimicking these effects when analyzing simulation data or developing experimental approaches that overcome these limitations.

We have drawn connections between this structural heterogeneity and protein function (discussed later) and developed simplified models that allow us to delve into the details of the physical chemistry underlying these fluctuations, discussed next.

DOI: 10.1021/ar500351b
Acc. Chem. Res. 2015, 48, 1098−1105

2.2. Utility of Side-Chain Fluctuations

The substantial side-chain fluctuations found in the work discussed above[5] and in prior simulations of natively folded proteins[3] invite us to consider the implications of such motions in terms of their potential contributions to biological function. Although a comprehensive picture of the utility of these native-state fluctuations is lacking, accumulating evidence from calorimetry and NMR experiments attests to the important roles both backbone and side-chain motions can play in regulating the thermodynamics of protein interactions.[12−15,31]

Inspired by these observations, we set out to explore the nature of side-chain fluctuations using a simple model of the natively folded protein, hoping to better understand the potential functionality of these smaller-scale yet ubiquitous motions. Briefly, our model consisted of an atomic-level representation of the protein with appropriate steric repulsions, van der Waals attractions, salt bridges, hydrogen bonds, and implicit solvent interactions.[4] We considered the equilibrium fluctuations associated with the side-chain torsional degrees of freedom and kept the backbone fixed in its crystalline configuration.[4] The simplicity of this model and our ability to sample through sterically inaccessible regions of conformational space (see ref 4), allowed us to make an exhaustive survey of side-chain variations that would not have been possible with more detailed models.

Using this approach, we probed the effect of different interatomic forces on the magnitude of side-chain fluctuations among a set of 12 small, single-domain proteins and examined their distribution within the proteins.[4] We calculated both the total entropy associated with these fluctuations (using a noninteracting reference system where all side chains were free to rotate) as well as the entropy associated with each side chain's occupancy of its different torsional wells. The former measure revealed the effect of different interactions on total side-chain torsional entropy, whereas the latter provided a way to assess the spatial distribution of fluctuations.

Our results were in agreement with those emerging from recent NMR,[32] crystallographic,[2] and computational[1,3,5,33] studies, indicating that significant side-chain fluctuations are present in natively folded proteins and, rather than just residing on the protein's surface, these motions are heterogeneously distributed throughout. In addition, by considering each interaction separately, we were able to examine in more detail how the various forces present in these proteins act to restrain the fluctuations present in the noninteracting system and thereby reduce the entropy, see Figure 2. Interestingly, we found that each kind of interaction affected the reduction in side-chain fluctuations similarly across the full set of 12 proteins: the presence of salt bridges and hydrogen bonds resulted in the greatest reduction of side-chain flexibility (and thus entropy) for most of the proteins, the presence of steric repulsions and the presence of van der Waals attractions also both resulted in a substantial reductions in fluctuations, and the reduction in side-chain entropy due to solvation was significantly less and acted mainly on the surface residues. It is important to note, however, that these simulations were done in implicit solvent, meaning that the effect of solvent fluctuations on protein motion was not explored.

Using the same model, we also set out to calculate the entropic contributions from these fluctuations to the binding free energy of calmodulin to four small peptide ligands that represent the binding domains of a few of its numerous binding partners. Isothermal calorimetry experiments by Wand and co-workers had identified this system as one in which the entropic contribution to the thermodynamics of binding helped tune the overall binding free energy to a similar level across the different peptides, despite disparate enthalpic contributions.[14] Further work in their group then showed a significant correlation between the calorimetrically determined entropy and the entropy contributions estimated from NMR relaxation measurements that report on the motions of methyl-bearing side chains.[14] Through our simulations of the system, we were able to calculate more comprehensive side-chain entropic terms that accounted for the motions of all side chains while also considering the correlations among them. These calculated side-chain entropies demonstrated a remarkable correlation to the isothermal calorimetry results, suggesting that nearly all of the variation observed in the entropy of binding between the different ligands could be attributed to variations in side-chain torsional fluctuations,[4] see Figure 3. Given calmodulin's multitude of binding partners,[14] having the ability to fine-tune its many interactions is crucial for it to properly execute its regulatory function; adjusting its side-chain fluctuations in response to binding appears to be one of the key tools it uses to do so.[4,14]

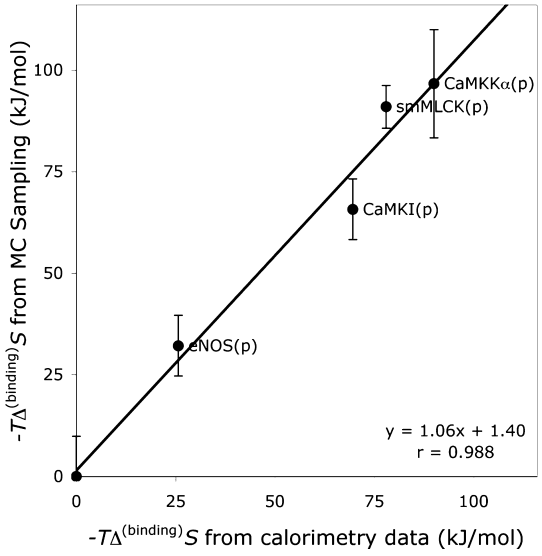


Figure 3. Entropic contributions to the binding free energies, $-T\Delta^{(\text{binding})}S$ for four CaM−peptide complexes. Results calculated from our simulations of side-chain fluctuations are plotted against the corresponding calorimetric measurements from ref 14. Unbound CaM is shown on the plot as the reference point at (0,0). Reproduced with permission from ref 4. Copyright 2009 Elsevier.

The temperature of the simulation also influenced the fluctuations we observed, in somewhat surprising ways. Although in general the fluctuations associated with each side chain are reduced upon cooling, the fluctuations of a few residues appear to *increase* upon cooling, most likely as a result of neighboring residues settling into a single rotameric well, leaving more space available.[6] This result is in agreement with observations from NMR and crystallographic studies[35−38] and confirms, in a more realistically complex system, a similar result that emerged from a small cluster model of interacting residues.[37] Such intricacies in the side chains' dynamic response to temperature change could prove important in stabilizing the fold and function of proteins that operate within extreme or highly variable environments.

Lastly, and perhaps most importantly, as part of our calculations, we were able to assess the range of side-chain

DOI: 10.1021/ar500351b
Acc. Chem. Res. 2015, 48, 1098−1105

torsional entropy values that we can expect to arise from typical variations in natively folded environments at physiological temperature. It turns out that this term is substantial: among our set of 12 small and well-folded proteins, we observed side-chain entropic contributions to the free energy that ranged over 0.6 kJ/(mol·300 K) per rotatable bond.[4] This value suggests that, for a protein with a small binding pocket consisting of 10 residues with an average of one or more rotatable bonds per residue, a change in the environment of that pocket due to ligand binding could readily result in a change of 6 kJ/mol to the free energy, more than double the value of k_BT at physiological temperature. Entropic contributions of this magnitude have clear implications for the stability of natively folded proteins as well as their ability to adjust their catalytic function in response to various environmental factors, including the binding of signaling molecules.

3. CORRELATED FLUCTUATIONS AND ALLOSTERIC REGULATION

The fluctuations discussed in the previous section take place within a very dense environment. Natively folded proteins are about as dense as organic crystals,[39] and the liquid-like fluctuations observed above in ref 5 occur without significant unfolding. These fluctuations must therefore arise from highly correlated rearrangements among the protein's constituent atoms, and recent work in NMR and crystallography have confirmed the presence of such dynamic networks.[40,41] Furthermore, when these correlated subsets span long distances within natively folded structures, they provide potential routes for the conveyance of allosteric information.

Although originally conceived of as a cooperative binding effect between identical subunits in multiunit protein complexes, such as hemoglobin,[11] the concept of allosteric regulation has evolved[42] and is now recognized as any modification of protein function at one location resulting from the perturbation of the protein at some other location. As such, allosteric regulation has been observed to function within a single domain protein[43] and across protein−protein interfaces,[44] through diverse mechanisms such as cryptic allostery,[45] where regulator molecules bind to small, transiently formed pockets, and dynamic allostery,[46] where the protein fluctuations themselves are modified.[22] In this section, we discuss our recent investigations into two ways in which correlations among the natively folded protein's fluctuations can mediate functional allostery: through the regulation of cryptic binding sites[7] and through the regulation of side-chain fluctuations.[8]

3.1. Cryptic Allostery

Cryptic allosteric sites are binding pockets that are not apparent from crystallographic structures yet can bind small molecules and communicate with distant sites. Such sites could serve as valuable drug targets, but they remain difficult to identify. Most of the known cryptic allosteric sites have been discovered by serendipity or high-throughput screens.[45,47,48] A site-directed method called tethering has also been of value[49] and could be even more powerful if combined with accurate predictions of where on a protein to search for such sites.

Inspired by the substantial fluctuations described above, we hypothesized that it may be possible to detect cryptic allosteric sites from a protein's dynamics, even in the absence of an allosteric modulator.[7] To test this hypothesis, we developed methods for detecting two signature structural fluctuations: (1) the formation of a pocket where a small molecule can potentially bind and (2) the presence of some form of correlated motions capable of communicating a binding event at one of these pockets to distant sites. Indeed, searching for these signatures readily identified a known hidden allosteric site and predicted the existence of a number of new sites (Figure 4). Experimental tests of these predictions are now ongoing. As before, we have also developed simplified models for dissecting how information flows within proteins.

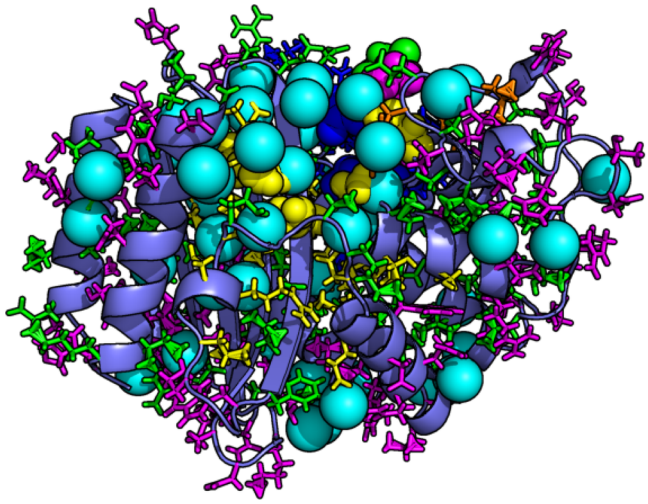


Figure 4. A multitude of potential cryptic allosteric sites in β-lactamase. The cyan spheres mark the locations of pockets with radii large enough to accommodate a few atoms of a small molecule (2−5 Å) that are present at least 1% of the time in our model. Side chains within each coupled community are rendered in the same color. They are depicted as spheres if they are in the active site and as sticks otherwise. There are five communities for β-lactamase. Almost every community contains at least one active site residue in each protein, so the vast majority of transient pockets could serve as cryptic sites, though multiple of these small pocket elements are required to accommodate a drug-like molecule. Reproduced with permission from ref 7. Copyright 2012 National Academy of Sciences.

3.2. Long-Range Correlations in Side-Chain Fluctuations

Using the same model described in section 2.2, we found that even just the fluctuations of side chains on a static backbone can give rise to long-range correlations.[8] By measuring the mutual information between pairs of residues, we quantified the degree of correlation between the motions of all side chains and probed how this correlation emerges from the various ways in which they interact with one another, through sterics, dispersion attractions, salt bridges, hydrogen bonds, and hydrophobic interactions. The results indicate that all interaction types give rise to some degree of correlation but that the longer-range correlations emerge when dispersion attractions, hydrogen bonding, and salt bridges are considered,[8] see Figure 5. These correlations are observable over the entire spatial extent of both proteins examined, reaching well beyond the range of any individual pairwise interaction present in the system. These results support the idea that dynamic coupling among residues is not just associated with known allosteric systems but is instead a general feature of folded proteins.[50] Interestingly, although it was long assumed that backbone motions must facilitate any correlated rearrangements, we showed in this work that they need not be present for significant correlations to arise among side-chain fluctuations.

Notably, recent work looking at correlated side-chain motions with a more detailed model suggests that backbone motions and

DOI: 10.1021/ar500351b
Acc. Chem. Res. 2015, 48, 1098−1105

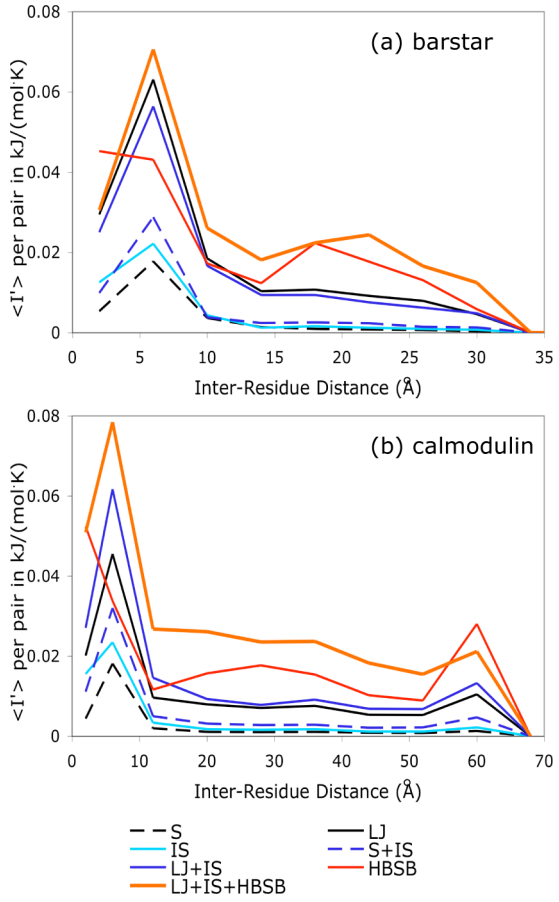


Figure 5. Long-range correlation among side-chain fluctuations as observed in simulations of such fluctuations within (a) barstar and (b) calmodulin. Correlation is quantified by the average excess mutual information, I', as a function of the distance between C_α atoms (see ref 8 for details). Results are shown for various atomic interactions: repulsive steric (S), implicit solvent (IS), Lennard-Jones (LJ), and hydrogen bonding (HB) and salt bridges (SB). Reproduced with permission from ref 8. Copyright 2011 Public Library of Science.

explicit solvent do not mitigate these correlations.[51] In this study, the authors also point out that, due to a protein's finite size, the number of residue pairs at long distance attenuates on the scale of the molecule's diameter. Viewed in aggregate, the predominant supply of correlation is therefore among side chains separated by modest distance.[51] As experiments become capable of probing in detail the subtler form of allostery that we have discussed, it will be interesting to see whether a natural length scale emerges from the competition between the greater availability of correlated residues closer to one another and the likely greater utility of communication between residues separated by longer distances.

In our work, we found that mutating single residues had measurable effects on the motions of other quite distant residues in both barstar and calmodulin.[8] This observation matches well with recent results from NMR and computational studies that emphasize the importance of correlated side-chain motions in allosterically regulating biological functions as diverse as iron uptake through pores in ferritin,[52] Nipah virus membrane fusion,[53] and interdomain signaling in the regulatory protein Pin1.[54] Overall, our results indicate that side-chain motions can be correlated over long distances in the folded protein, even in the absence of backbone motion.[8] These correlations span the full range of the protein and arise from several types of interactions, leading us to conclude that they are a robust feature of the natively folded protein and can readily contribute to the regulation of protein interactions.

4. CONCLUSION

Our computational studies buttress a growing body of experimental work that increasingly supports the idea of proteins as highly dynamic entities. In addition to quantitative agreement with experiments, our methods provide a uniquely detailed picture of the fluctuations that folded proteins undergo; using this approach, we are able to make perturbations that would be impossible experimentally, such as turning off electrostatic interactions, and thereby gain insight into the physical basis of this conformational heterogeneity.

Beyond quantifying the structural heterogeneity of various proteins, we have begun to tie their fluctuations to vital aspects of protein function. Our work highlights the importance of these seemingly small fluctuations for protein interactions and regulation. We observe that alterations in even just the side-chain fluctuations of a protein can influence its thermodynamics of binding dramatically.[4] We also find that the correlation among such motions provides mechanisms for allosteric regulation.[7,8] Indeed, the high degree of correlated motions found within folded proteins, even those not associated with any known allosteric function,[8,50] and the robust way in which these correlations arise, together suggest that biology has made use of and refined existing networks of correlated fluctuations as needed to aid in the allosteric regulation of protein function.[8]

Such networks are exceedingly difficult to probe experimentally, necessitating the use of computational models, such as those described here. However, the agreement between experimental and computational measures of fluctuation and correlation must be further improved. It is expected that, in the future, more refined force fields and better sampling algorithms will provide a more accurate match to *in vivo* systems. At the same time, the interpretation of various experimental measurements and their implication for protein motion are also expected to improve.

The observations outlined here suggest novel ways for us to intervene in the regulation of various disease-implicated catalytic proteins. In the past, entropic and allosteric considerations have been largely absent from docking and drug design strategies. However, with a better understanding of the native-state fluctuations, we may be able to rationally design improvements to the function of catalytic proteins by fine-tuning entropic contributions to their binding free energies or by utilizing pre-existing dynamic networks to alter their function allosterically.

The picture that has emerged from this work depicts the folded protein as a stable but dynamic structure riddled with networks of correlated motions. These dynamic features have been employed to aid the protein's functional response to environmental changes, in ways that enhance both its stability, as in calmodulin maintaining its function in the presence of a variety of binding partners,[55] and its adaptability, as in β-lactamase altering its function in the presence of an allosteric regulator.[7] A necessary consequence of the forces at work within these biomolecules, small-scale fluctuations and the correlations among them, appear crucial to the reliability and environmental-responsivity of folded proteins.

DOI: 10.1021/ar500351b
Acc. Chem. Res. 2015, 48, 1098−1105

AUTHOR INFORMATION

Corresponding Author

*E-mail: geissler@berkeley.edu.

Funding

G.R.B. holds a Career Award at the Scientific Interface from the Burroughs Wellcome Fund and was also supported by the Miller Institute for Basic Research in Science. P.L.G. was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, through the Materials Sciences Division (MSD) of the Lawrence Berkeley National Laboratory (LBNL), under Contract DE-AC02-05CH11231.

Notes

The authors declare no competing financial interest.

Biographies

Kateri H. DuBay was born October 13, 1980, in Atlanta, Georgia. She received her B.S. in Biochemistry at Georgetown University (2002), her M.Phil. in Chemistry at the University of Cambridge (2004), and her Ph.D. in Chemistry at UC Berkeley (2009). She was a Postdoctoral Scholar at UC Berkeley (2009−2010) and a Postdoctoral Research Scientist at Columbia University (2010−2013) before becoming Assistant Professor of Chemistry at the University of Virginia (2014). Her research efforts to date have focused on the functional implications of variations in primary structure and fluctuations within conjugated polymers and proteins.

Greg R. Bowman was born February 24, 1984, in Frankfurt, Germany. He received his B.S. in computer science from Cornell University (2006) and his Ph.D. in biophysics from Stanford University (2010), and he was a Postdoctoral Fellow at the University of California, Berkeley, from 2011 to 2014. He is now Assistant Professor of Biochemistry and Molecular Biophysics at Washington University School of Medicine (2014 to the present). Greg has been an important player in the development of Markov state model methods for simulating long time scale dynamics in atomistic detail. A major focal point of his recent research has been understanding the alternative structures that folded proteins adopt and the opportunities these structures present for manipulating protein function, for example, by finding unexpected allosteric sites.

Phillip L. Geissler was born March 27, 1974, in Ithaca, NY. He received his B.A. in Chemistry at Cornell University and his Ph.D. in Chemistry at UC Berkeley, and he was a Science Fellow at MIT from 2001 to 2003. Since then he has been Assistant, Associate, and Full Professor of Chemistry at UC Berkeley. His research interests include fluctuations in complex molecular systems, with current focuses on dynamics of self-assembly, solvation near soft interfaces, and functional implications of structural variability in biological systems on scales ranging from molecular to cellular.

REFERENCES

(1) Lindorff-Larsen, K.; Best, R. B.; Depristo, M. A.; Dobson, C. M.; Vendruscolo, M. Simultaneous determination of protein structure and dynamics. *Nature* **2005**, *433*, 128−132.

(2) Fraser, J. S.; Clarkson, M. W.; Degnan, S. C.; Erion, R.; Kern, D.; Alber, T. Hidden alternative structures of proline isomerase essential for catalysis. *Nature* **2009**, *462*, 669−673.

(3) Kussell, E.; Shimada, J.; Shakhnovich, E. I. Excluded volume in protein side-chain packing. *J. Mol. Biol.* **2001**, *311*, 183−193.

(4) DuBay, K. H.; Geissler, P. L. Calculation of proteins' total side-chain torsional entropy and its influence on protein-ligand interactions. *J. Mol. Biol.* **2009**, *391*, 484−497.

(5) Bowman, G. R.; Geissler, P. L. Extensive conformational heterogeneity within protein cores. *J. Phys. Chem. B* **2014**, *118*, 6417−6423.

(6) DuBay, K. H. Protein side-chain fluctuations: A computational study of entropy and correlation, Ph.D. Thesis, University of California, Berkeley, 2009.

(7) Bowman, G. R.; Geissler, P. L. Equilibrium fluctuations of a single folded protein reveal a multitude of potential cryptic allosteric sites. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 11681−11686.

(8) Dubay, K. H.; Bothma, J. P.; Geissler, P. L. Long-range intra-protein communication can be transmitted by correlated side-chain fluctuations alone. *PLoS Comput. Biol.* **2011**, *7*, No. e1002168.

(9) Noji, H.; Yasuda, R.; Yoshida, M.; Kinosita, K., Jr. Direct observation of the rotation of F1-ATPase. *Nature* **1997**, *386*, 299−302.

(10) Svoboda, K.; Schmidt, C. F.; Schnapp, B. J.; Block, S. M. Direct observation of kinesin stepping by optical trapping interferometry. *Nature* **1993**, *365*, 721−727.

(11) Perutz, M. F. Stereochemistry of cooperative effects in haemoglobin. *Nature* **1970**, *228*, 726−739.

(12) Zidek, L.; Novotny, M. V.; Stone, M. J. Increased protein backbone conformational entropy upon hydrophobic ligand binding. *Nat. Struct. Biol.* **1999**, *6*, 1118−1121.

(13) Bernini, A.; Ciutti, A.; Spiga, O.; Scarselli, M.; Klein, S.; Vannetti, S.; Bracci, L.; Lozzi, L.; Lelli, B.; Falciani, C.; Neri, P.; Niccolai, N. NMR and MD studies on the interaction between ligand peptides and alpha-bungarotoxin. *J. Mol. Biol.* **2004**, *339*, 1169−1177.

(14) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. Conformational entropy in molecular recognition by proteins. *Nature* **2007**, *448*, 325−329.

(15) Arumugam, S.; Gao, G.; Patton, B. L.; Semenchenko, V.; Brew, K.; Doren, S. R. V. Increased backbone mobility in beta-barrel enhances entropy gain driving binding of N-TIMP-1 to MMP-3. *J. Mol. Biol.* **2003**, *327*, 719−734.

(16) Tzeng, S.-R.; Kalodimos, C. G. Protein activity regulation by conformational entropy. *Nature* **2012**, *488*, 236−240.

(17) Wuthrich, K.; Wagner, G. Internal motion in globular proteins. *Trends Biochem. Sci.* **1978**, *3*, 227−230.

(18) Englander, S. W.; Downer, N. W.; Teitelbaum, H. Hydrogen exchange. *Annu. Rev. Biochem.* **1972**, *41*, 903−924.

(19) McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of folded proteins. *Nature* **1977**, *267*, 585−590.

(20) Levitt, M.; Huber, R. Molecular dynamics of native protein. II. Analysis and nature of motion. *J. Mol. Biol.* **1983**, *168*, 621−657.

(21) Brooks, C. L. I.; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure, And Thermodynamics*; Wiley Press: New York, 1988; Vol. *71*.

(22) Cooper, A.; Dryden, D. T. Allostery without conformational change. A plausible model. *Eur. Biophys. J.* **1984**, *11*, 103−109.

(23) Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hubner, C. G.; Kern, D. Intrinsic motions along an enzymatic reaction trajectory. *Nature* **2007**, *450*, 838−844.

(24) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **2006**, *313*, 1638−1642.

(25) Hilser, V. J.; Dowdy, D.; Oas, T. G.; Freire, E. The structural distribution of cooperative interactions in proteins: Analysis of the native state ensemble. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 9903−9908.

(26) Perryman, A. L.; Lin, J. H.; McCammon, J. A. HIV-1 protease molecular dynamics of a wild-type and of the V82F/I84V mutant: Possible contributions to drug resistance and a potential new target site for drugs. *Protein Sci.* **2004**, *13*, 1108−1123.

(27) McClendon, C. L.; Friedland, G.; Mobley, D. L.; Amirkhani, H.; Jacobson, M. P. Quantifying correlations between allosteric sites in thermodynamic ensembles. *J. Chem. Theory Comput.* **2009**, *5*, 2486−2502.

(28) Schwartz, S. D.; Schramm, V. L. Enzymatic transition states and dynamic motion in barrier crossing. *Nat. Chem. Biol.* **2009**, *5*, 551−558.

DOI: 10.1021/ar500351b
Acc. Chem. Res. 2015, 48, 1098−1105

(29) Bowman, G. R.; Huang, X.; Pande, V. S. Network models for molecular kinetics and their initial applications to human health. *Cell Res.* **2010**, *20*, 622−630.
(30) Chodera, J. D.; Noe, F. Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.* **2014**, *25*, 135−144.
(31) Wand, A. J. The dark energy of proteins comes to light: conformational entropy and its role in protein function revealed by NMR relaxation. *Curr. Opin. Struct. Biol.* **2013**, *23*, 75−81.
(32) Igumenova, T. I.; Frederick, K. K.; Wand, A. J. Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution. *Chem. Rev.* **2006**, *106*, 1672−1699.
(33) Baxa, M. C.; Haddadian, E. J.; Jumper, J. M.; Freed, K. F.; Sosnick, T. R. Loss of conformational entropy in protein folding calculated using realistic ensembles and its implications for NMR-based calculations. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 15396−15401.
(34) Brudler, R.; Meyer, T. E.; Genick, U. K.; Devanathan, S.; Woo, T. T.; Millar, D. P.; Gerwert, K.; Cusanovich, M. A.; Tollin, G.; Getzoff, E. D. Coupling of hydrogen bonding to chromophore conformation and function in photoactive yellow protein. *Biochemistry* **2000**, *39*, 13478−13486.
(35) Lee, A. L.; Wand, A. J. Microscopic origins of entropy, heat capacity and the glass transition in proteins. *Nature* **2001**, *411*, 501−504.
(36) Song, X.; Flynn, P. F.; Sharp, K. A.; Wand, A. J. Temperature dependence of fast dynamics in proteins. *Biophys. J.* **2007**, *92*, L43−L45.
(37) Lee, A. L.; Sharp, K. A.; Kranz, J. K.; Song, X.-J.; Wand, A. J. Temperature dependence of the internal dynamics of a calmodulin-peptide complex. *Biochemistry* **2002**, *41*, 13814−13825.
(38) Fraser, J. S.; van den Bedem, H.; Samelson, A. J.; Lang, P. T.; Holton, J. M.; Echols, N.; Alber, T. Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 16247−16252.
(39) Chothia, C. Structural invariants in protein folding. *Nature* **1975**, *254*, 304−308.
(40) Clarkson, M. W.; Lee, A. L. Long-range dynamic effects of point mutations propagate through side chains in the serine protease inhibitor eglin c. *Biochemistry* **2004**, *43*, 12448−12458.
(41) van den Bedem, H.; Bhabha, G.; Yang, K.; Wright, P. E.; Fraser, J. S. Automated identification of functional dynamic contact networks from X-ray crystallography. *Nat. Methods* **2013**, *10*, 896−902.
(42) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. The ensemble nature of allostery. *Nature* **2014**, *508*, 331−339.
(43) Volkman, B. F.; Lipson, D.; Wemmer, D. E.; Kern, D. Two-state allosteric behavior in a single-domain signaling protein. *Science* **2001**, *291*, 2429−2433.
(44) Szilagyi, A.; Nussinov, R.; Csermely, P. Allo-network drugs: Extension of the allosteric drug concept to protein- protein interaction and signaling networks. *Curr. Top. Med. Chem.* **2013**, *13*, 64−77.
(45) Horn, J. R.; Shoichet, B. K. Allosteric inhibition through core disruption. *J. Mol. Biol.* **2004**, *336*, 1283−1291.
(46) Popovych, N.; Sun, S.; Ebright, R. H.; Kalodimos, C. G. Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* **2006**, *13*, 831−838.
(47) Lundqvist, T.; Fisher, S. L.; Kern, G.; Folmer, R. H.; Xue, Y.; Newton, D. T.; Keating, T. A.; Alm, R. A.; de Jonge, B. L. Exploitation of structural and regulatory diversity in glutamate racemases. *Nature* **2007**, *447*, 817−822.
(48) Ceccarelli, D. F.; et al. An allosteric inhibitor of the human Cdc34 ubiquitin-conjugating enzyme. *Cell* **2011**, *145*, 1075−1087.
(49) Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A. Site-directed ligand discovery. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9367−9372.
(50) Clarkson, M. W.; Gilmore, S. A.; Edgell, M. H.; Lee, A. L. Dynamic coupling and allosteric behavior in a nonallosteric protein. *Biochemistry* **2006**, *45*, 7693−7699.
(51) Wu, S.; Jun Lee, C.; Pedersen, L. G. Analysis on long-range residue-residue communication using molecular dynamics. *Proteins* **2014**, *82*, 2896−901.
(52) Ruvinsky, A. M.; Vakser, I. A.; Rivera, M. Local packing modulates diversity of iron pathways and cooperative behavior in eukaryotic and prokaryotic ferritins. *J. Chem. Phys.* **2014**, *140*, No. 115104.
(53) Varma, S.; Botlani, M.; Leighty, R. E. Discerning intersecting fusion-activation pathways in the Nipah virus using machine learning. *Proteins* **2014**, *82*, 3241−3254.
(54) Wilson, K. A.; Bouchard, J. J.; Peng, J. W. Interdomain interactions support interdomain communication in human Pin1. *Biochemistry* **2013**, *52*, 6968−6981.
(55) Frederick, K. K.; Kranz, J. K.; Wand, A. J. Characterization of the backbone and side chain dynamics of the CaM-CaMKIp complex reveals microscopic contributions to protein conformational entropy. *Biochemistry* **2006**, *45*, 9841−9848.

DOI: 10.1021/ar500351b
Acc. Chem. Res. 2015, 48, 1098−1105