On the role of frustration in the energy landscapes of allosteric proteins

Diego U. Ferreiro[a], Joseph A. Hegler[b,c], Elizabeth A. Komives[b], and Peter G. Wolynes[b,c,1]

[a]Protein Physiology Lab, Department of Biological Chemistry, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina, Buenos Aires, Argentina C1428EGA; [b]Department of Chemistry and Biochemistry; and [c]Center for Theoretical Biological Physics, University of California at San Diego, La Jolla, CA 92107.

Contributed by Peter G. Wolynes, December 16, 2010 (sent for review November 24, 2010)

Natural protein domains must be sufficiently stable to fold but often need to be locally unstable to function. Overall, strong energetic conflicts are minimized in native states satisfying the principle of minimal frustration. Local violations of this principle open up possibilities to form the complex multifunnel energy landscapes needed for large-scale conformational changes. We survey the local frustration patterns of allosteric domains and show that the regions that reconfigure are often enriched in patches of highly frustrated interactions, consistent both with the idea that these locally frustrated regions may act as specific hinges or that proteins may "crack" in these locations. On the other hand, the symmetry of multimeric protein assemblies allows near degeneracy by reconfiguring while maintaining minimally frustrated interactions. We also anecdotally examine some specific examples of complex conformational changes and speculate on the role of frustration in the kinetics of allosteric change.

minimal frustration principle | protein folding | protein function

Allostery and large-scale conformational changes are widespread in molecular biology but historically have been considered to be exceptional and somewhat mysterious. In fact, cryobiochemical (1) and single molecule (2) experiments show that the underlying energy landscapes of all biomolecules are generally quite complex (3). These findings surprised many because there has been so much success in modeling even large pieces of biological machinery as simple chemical entities obeying elementary laws of equilibrium and kinetics (4). The mystery of allostery was thus a first hint of landscape complexity (5). In contrast to experimentalists, who were surprised by emergent complexity, theorists are more surprised by the seeming simplicity of the free energy landscape of proteins at physiological temperatures. Theorists expect that the apparent randomness of a protein sequence will result in many competing forces between residues, and thus structurally disparate states should be at least transiently populated (6). Indeed statistical mechanical theory suggests completely random heteropolymers have rugged landscapes, like glasses, which provide paradigms of complex kinetics (7–9). The resolution of this dialectic lies in evolution: Proteins emerge from selected sequences that give rise to organized energy landscapes. Most of this organization encodes the ability of the molecule to spontaneously find a fairly specific (although decidedly not unique) configuration, the so-called folded or average native structure. By having specific structures, proteins become limited in their range of interaction partners thus allowing complex networks of biological interactions to be built up. Overall, the energy landscapes of proteins resemble a rough funnel leading toward the native state (10, 11). This funnel structure is only possible for those selected sequences that are chosen so that energetic conflicts are for the most part avoided and the native structure is more stable than expected for random associating residues. This thesis is known as the "principle of minimal frustration" (6). If the minimal frustration principle is satisfied everywhere, the protein molecule becomes a beautiful sculpture with a tendency to remain intact and move as a rigid body. We will explore here how local violations of the minimal frustration principle open up possibilities for more complex energy landscapes needed for allostery and large-scale conformational changes (12, 13).

Multiple funnels to structurally distinct low-free-energy states can also be achieved by other mechanisms (14), symmetry being the main route to such degeneracy (15). Nearly rigid macromolecular subunits can pack in a number of symmetry equivalent ways with similar free energies. This mathematical phenomenon is the core of the brilliant insight of Monod et al. (16) which spotlighted symmetric multimeric proteins as prime candidates for allostery.

To investigate the role of frustration in large-scale conformational changes, we need to locate sites both where the minimal frustration principle is strongly violated as well as the web of minimally frustrated interactions that impart rigidity to much of the protein structure. A simple heuristic based on energy landscape theory proposed by us earlier is able to do this localization. We previously showed that some clusters of highly frustrated interactions signal binding sites for protein–protein assembly and recognition (17). Using the same algorithm, we now survey in this paper a large number of proteins that undergo large-scale configurational motions, generally as monomers. Our survey sample consists of proteins crystallized in pairs of alternative forms (18). Usually the alternate structures are stabilized by adding appropriate ligands, although in some cases, modest sequences changes have allowed the capture in the crystal of the alternate configurations. We show that the more rigid parts of the proteins, which are locally structurally superimposable, are connected by a dense web of largely minimally frustrated interactions. On the other hand, regions that are highly frustrated often reconfigure locally between the two forms. In some cases, frustrated regions display rather extensive reconfigurations of compact regions. In other cases, the frustrated clusters localize around apparent pivot points between the more rigid elements. The accompanying structural change is then a combination of local rearrangements in the vicinity of these pivots and large-amplitude rigid-body motions through space of the minimally frustrated elements neighboring them. The motion thus superficially resembles a macroscopic hinge. The frustration analysis is consistent both with the idea that these locally frustrated regions may be specific hinges or that the proteins crack in these locations (12). Cracking is encouraged by the low local stability that accompanies high local frustration (17, 19).

We first briefly review the algorithm for localizing frustration and the criteria we use for locating residues locally displaced between the pairs of proteins. We then show that frustrated clusters are indeed colocated with those residues whose local

Author contributions: D.U.F., J.A.H., E.A.K., and P.G.W. designed research; D.U.F. and J.A.H. performed research; D.U.F., E.A.K., and P.G.W. analyzed data; and D.U.F., E.A.K., and P.G.W. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: pwolynes@ucsd.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1018980108/-/DCSupplemental.

environment shifts between the two structures. The survey provides examples in which only one structure of a pair has a relevant frustrated cluster whose frustration disappears in the other conformation. Other times, both structures are frustrated in the same place but different specific frustrated interacting residues are interchanged with each other. There are no examples of structure pairs where the locally displaced regions were completely minimally frustrated in both configurations for the monomeric allostery, but this phenomenon does happen in multimers and indeed occurs for the venerable example of haemoglobin. We then anecdotally examine some more complex conformational changes such as occur in Adenylate kinase where there are multiple loci of rearrangement. The catabolite activator protein (CAP) protein has been suggested as an example of "dynamic" allostery: No distinct mean structural changes have been found but changes in the local dynamics are manifested in NMR relaxation experiments (20). Finally we comment on the role of frustration in the kinetics of allosteric change and its role in cracking.

Results and Discussion

We use a curated database of pairs of homologous proteins whose structure has been solved in at least two different states (18). We calculate the local frustration indices for each protein using a local version of the global gap criterion for a protein to be minimally frustrated (17). The energy function used is a low-resolution nonadditive water-mediated potential which is transferable and quite successful in ab initio protein structure prediction (21). Briefly, to compute the local frustration indices, every contacting residue pair in an input structure is exhaustively mutated and the total contact energy of the protein is calculated using the associative memory Hamiltonian optimized with water-mediated interactions force field (22). The native energy is compared with the distribution of decoy energies using a Z-score criterion, defining a "frustration index" (17). This Z-score criterion, when used for the entire protein, delimits minimally frustrated sequences from those that exhibit glassy dynamics. Two complementary ways of constructing the decoy set allow us to compare interactions in the native structure to other possible sets of interactions. From a decoy set comprising structurally conservative mutants, native interactions of a residue are compared to those present when other amino acid residues are placed at the same position, yielding a "mutational frustration index." From a decoy set comprising alternative but still compact structures of the native sequence, a

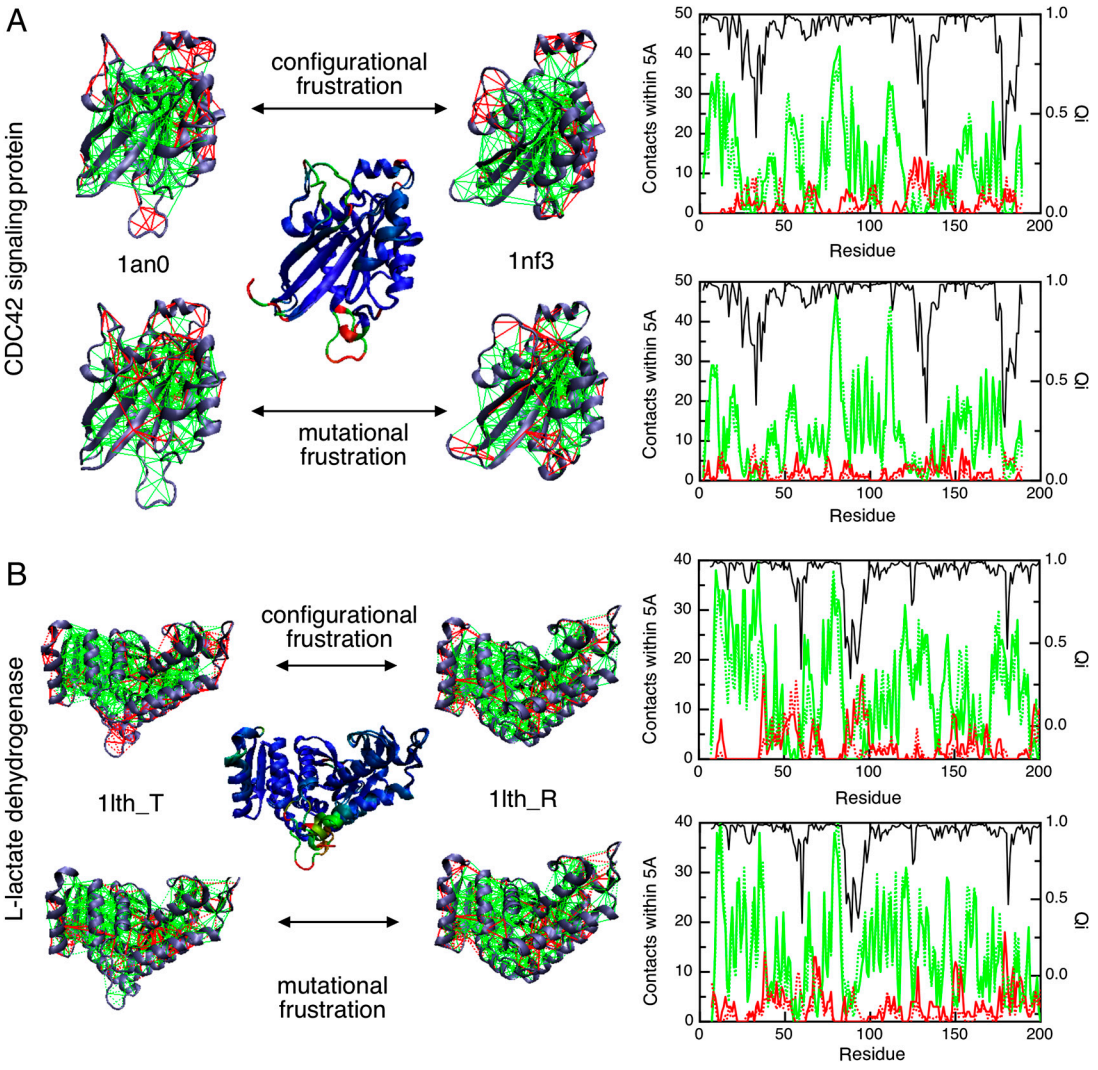

Fig. 1. Gallery of the localized frustration and minimally frustrated networks in allosteric proteins. A structural alignment of both experimentally determined conformations is shown at the center, colored according to the structural deviation (blue low, red high). The individual conformations are shown at the sides. The protein backbone is displayed as ribbons, the interresidue interactions with solid lines. Minimally frustrated interactions are shown in green, highly frustrated interactions in red, neutral contacts are not drawn. At right, a quantification of the minimally frustrated interactions (green) or highly frustrated interactions (red) in the vicinity of each residue, in form A (solid, 1AN0 and 1LTH_T) or form B (dashed, 1NF3 and 1LTH_R). The local Q_i of each residue is shown in black.

residue's interactions in the native structure can be compared to those present in these alternative structures, yielding a "configurational frustration index" (17). A sample of the frustratograms used (total of 23 molecular pairs) can be found in Fig. S1.

We first analyze the distribution of frustration indices, without discriminating between pairs of related structures. Using the previously estimated cutoff values of the frustration indices (17), we found that allosteric protein domains are strongly connected by a web of minimally frustrated interactions encompassing about 40% of the total contacts. Only about 10% of the interactions can be considered to be "highly frustrated" and the remainder are "neutral." The distribution of frustration index in allosteric proteins is very close to the previously reported distribution in a random sample of small monomeric domains (17), supporting the notion that all natural monomeric proteins may potentially be allosteric (5, 23).

Examples of frustration patterns and allosterism are shown in Fig. 1. Pictured first are the frustratograms for CDC42 signaling protein. Both structural forms contain an extensive network of minimally frustrated interactions shown in green. This network holds together the more rigid core. Extended patches of frustrated residues as measured by the configurational frustration index are found at residues 130–150 and 180–190 of the 1AN0 form. In the alternate 1NF3 form, the 180–190 patch of frustrated residues become considerably less frustrated and is the region of the largest structural change. Interestingly, the loop in 1AN0, although highly frustrated as indicated by the configurational measure, is not highly frustrated when the mutational index is used, indicating that this region is poised to change structure via burial. The somewhat less extensive structural change of the frustrated patch at the 130–150 region of the molecule remains frustrated in both forms but exchanges specific instances of frustrated interactions.

A related pattern appears in the L-lactate dehydrogenase. This protein also has several patches that correspond with the binding sites to the partner along with a frustrated patch in the residues 80–100 of 1LTH_T, which reconfigures and considerably lowers its frustration level in the 1LTH_R form (Fig. 1*B*). About 10% of the contacts in that region change the frustration index from highly frustrated to minimally frustrated.

In order to quantify the nature of structural conservation between conformational substates and local frustration, we introduce a "local-Q" scoring parameter to locate the structural differences between protein pairs. We define Q_i as

$$Q_i = \frac{1}{(N-3)} \sum_{i \neq j}^{N} e^{-\left(\frac{r_A^{ij} - r_B^{ij}}{\sigma^2}\right)} \qquad \sigma = |i-j|^{0.15} \qquad \begin{matrix} |i-j| > 1 \\ \min(r_A^{ij} | r_B^{ij}) < 6 \ \text{Å} \end{matrix}$$

where r_A^{ij} and r_B^{ij} is the distance between the Cα of residues i and j in the A and B forms, respectively. The sum is over those residues closer than 6 Å in either form. Given two structures, this parameter quantifies their local structural deviations without making structural superpositions. It correlates with other deviation measures introduced by Daily and Gray (18) (Fig. S2). We classify residues as locally displaced or "mobile" if $Q_i < 0.9$ and as undisplaced or "static" if $Q_i > 0.9$.

Visual inspection suggests that locally displaced regions are enriched in patches of highly frustrated interactions (Fig. 1 and Fig. S1). To quantify this correlation, we compute the pair-distribution function between residues classified by displacement and the contacts in different frustration classes (Fig. 2). We found that regions with mobile residues are enriched in highly frustrated interactions up to 5*A* from the Cα (Fig. 2*A*). On the other hand, "undisplaced" residues are more closely correlated at long distances by minimally frustrated interactions than the overall topology of the contact network (Fig. 2*B*). Interactions classified with neutral frustration are homogeneously distributed in the contact network. The signals show up more strongly for the configurational frustration index (Fig. S3). This observation is consistent with other definitions of mobile or static residues (Fig. S4).

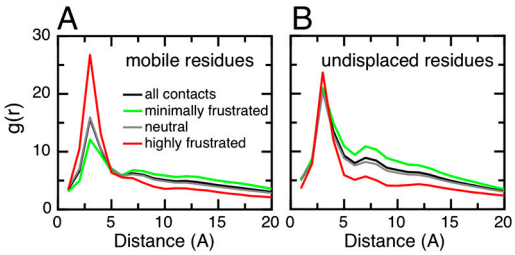

Fig. 2. Local frustration and residue displacement. The pair-distribution functions between the Cα of the residues classified by displacement and the center of mass of the contacts in different frustration classes was computed. The distributions for all contacts (black), minimally frustrated (green), neutral (gray), or highly frustrated contacts (red), are shown for the mobile (*A*) or undisplaced (*B*) residues.

The major structural "core" is conserved between the two conformational substates. Structural differences tend to localize on the protein surface (18) and often include the binding site of an allosteric effector (interactions with the allosteric effector are not included in calculating the frustration index). About 80% of the pairwise interactions are found in both substates, whereas 10% are exclusively found in one conformational substate or the other. Contacts common to both structures have similar frustration index distributions to those for contacts exclusively found in only one substate (Fig. 3). The common interactions do not dramatically change frustration class (Fig. S5), instead new contacts are formed in one substate that does not exist in the other, with similar local frustration distributions. Thus, generally, there is a balanced exchange of local frustration upon switching substates.

Allostery was first described as the regulation of protein activity through changes in quaternary structure of multisubunit complexes (16). In principle, preserving symmetry is an easy path for achieving low energy near degenerate conformation. Using the same counting argument that makes symmetric structures individually more easily designable (15), pairs of symmetric structures provide a statistically favored way of achieving a two-state-like switching behavior. We calculated the local frustration distribution in the multimeric forms of the allosteric proteins that are known (or predicted to) form quaternary architectures. On average, interfacial (quaternary) interactions prove to be less frustrated than monomeric (tertiary) interactions (Fig. 4). The interfaces are depleted in highly frustrated interactions, suggesting that symmetry may be reflected in the local frustration patterns when alternate configurations are used as decoys. On the other hand, the interfaces have a sharper distribution of frustration than monomers when the mutational frustration index is evaluated (Fig. 4*B*). The effect that mutations exert on quaternary

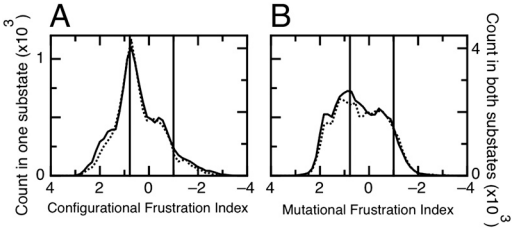

Fig. 3. Local frustration in allosteric pairs. The distribution of the configurational (*A*) and mutational (*B*) frustration indices were calculated for the contacts conserved between substates (solid) or exclusive to one substate (dashed). The vertical lines indicate the cutoff used to define minimally frustrated, neutral, or highly frustrated interactions.

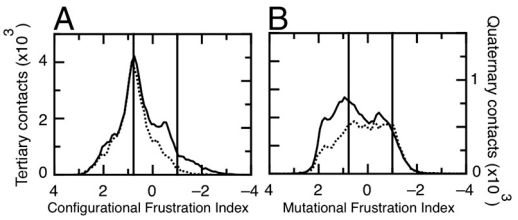

Fig. 4. Local frustration in oligomeric proteins. The distribution of the configurational (*A*) and mutational (*B*) frustration indices were calculated for the contacts internal to each monomer (tertiary, solid lines) or in between monomers (quaternary, dashed lines). The vertical lines indicate the cutoff used.

interactions is thus more neutral than the effect on the internal tertiary interactions of a subunit.

Several specific allosteric cases not included in the survey dataset are worth noting. One of these is the primal example of allosterism, haemoglobin. Here the interface between the subunits is not configurationally frustrated in either form. Rather a rigid-body rotation allows symmetrically related interactions involving a different knob-in-hole packing of the helices of one monomer on the other. Both packings are configurationally minimally frustrated (Fig. 5*C*).

Adenylate kinase also provides a celebrated example of large-scale conformational change (Fig. 5*A*). The opening and closing of the protein involves at least two reaction coordinates. Steered molecular dynamic calculations have identified several locations where "hinge" motions are thought to occur (24). Examining the "open" and "closed" frustratograms show the hinges are located in highly frustrated regions. Simultaneously, in the closed form, a more extensive minimally frustrated network of contacts rigidifies the molecule. Rigid-body motions of the lid and the core are involved in the closure. Miyashita et al. have examined motion of adenylate kinase along the low-frequency normal modes contributing to this closure and concluded that there were regions of high stress at the transition state for configurational change (12). This high-stress region can however crack or locally unfold releasing the strain and catalyzing the motion (25–27). This high-strain region is indeed highly frustrated in both forms (Fig. 5*A*).

The frustratogram of the dimeric form of CAP is shown as a last example (Fig. 5*B*). In this case, although there is negative cooperativity upon binding cAMP, there is no evidence for a change in average structure, but rather cAMP does cause a change of NMR parameters (20). The comparison of the exchange parameters along the sequence and the local frustration is shown in Fig. 5*B*. The dynamically perturbed region corresponds with the frustrated interactions at the dimer interface. Presumably frustration in this part of the molecule allows the existence of many configurational substates whose distribution changes upon substrate binding.

Concluding Remarks

The notion of frustration originated in the study of magnetic models, where at the local level it signaled a degeneracy of states (28). In an extended system, this degeneracy leads to glassy behavior. For proteins, most of the interactions are not frustrated (6), but evolution has made use of frustration in discrete parts of the protein allowing competing low-free-energy structures that can be differentially favored under differing thermodynamic conditions or when ligands bind (17). Thus localized frustration is a natural origin for allosteric transitions. The present structural

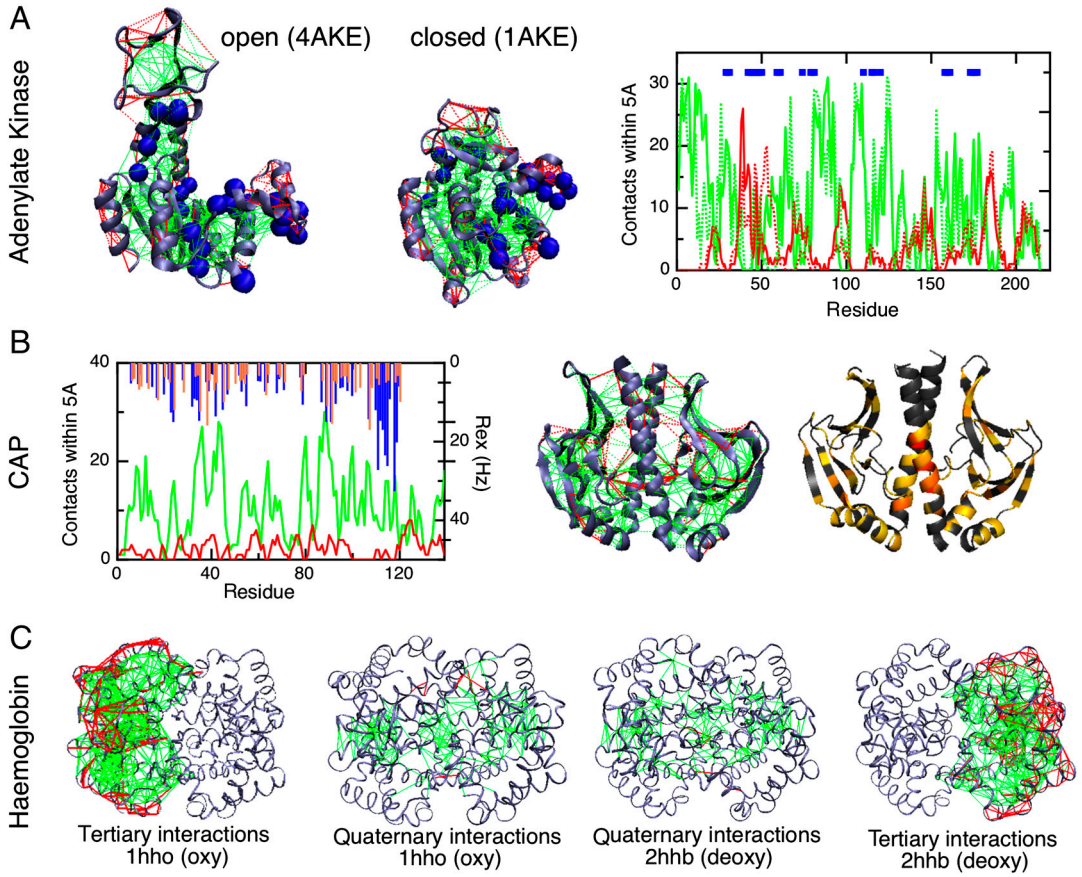

Fig. 5. Local frustration in paradigmatic examples. Frustratographs for the open and closed forms of Adenylate kinase are shown in *A*. The Cα of residues where hinge motions are thought to occur (24) are highlighted in blue. The local frustration pattern of CAP protein is shown in *B*, side to the change in NMR exchange parameter (20). The tertiary and quaternary frustration patterns of the oxy and deoxy forms of haemoglobin are shown in *C*.

survey shows that localized frustration is connected with local conformational flexibility and large-scale conformational changes.

The precise way in which localized frustration becomes associated with allostery requires further investigation (29). Frustration may simply allow a small discrete set of configurations involving local motions of the frustrated residues. The allosteric dynamical transition may then involve defined paths of transformation like a small molecule, or indeed like macroscopic machines with well-oiled hinges. Alternatively, the frustration may destabilize a part of the protein in favor of an ensemble of rather high entropy—in other words, a local region can unfold or crack. That possibility has been entertained previously (12) and nicely explains observations of denaturant catalyzed conformational change (30). Clearly resolving experimentally the issue of hinges versus cracks requires an analysis of the effect that local mutations on the interconformational kinetics using protein engineering methodology. Such analysis has its in silico counterpart in our frustration analysis. This question can also be addressed with simulations based on energy functions that capture the physics of frustration in proteins (29, 31).

Evolutionarily engineered frustration is not essential for allostery. The classic Wyman-Monod view of symmetry of multimeric assemblies leading to near degeneracy does not depend on frustration at all. Finally, because we are sometimes concerned with only a few k_bT of free energy, weakly assembled complexes may accidentally exhibit allostery without any very strong frustration, as suggested by Kuriyan and Eisenberg (32). Cooptation of such accidental happenings is not an uncommon biological trick (33).

Methods

We calculate the local frustration indices as described previously (17) using a curated database of allosteric proteins (18). More information can be found in *SI Methods*.

ACKNOWLEDGMENTS. We thank Babis Kalodimos for providing NMR data for CAP. This work was supported by National Institutes of Health Grants P01GM071862 and R03 TW008232-01 and National Science Foundation Grants PHY0216576 and PHY0225630 to the Center for Theoretical Biological Physics.

1. Austin RH, et al. (1975) Dynamics of ligand binding to myoglobin. *Biochemistry* 14:5355–5373.
2. Min W, et al. (2005) Fluctuating enzymes: Lessons from single-molecule studies. *Acc Chem Res* 38:923–931.
3. Frauenfelder H, Sligar SG, Wolynes PG (1991) The energy landscapes and motions of proteins. *Science* 254:1598–1603.
4. Fersht AR (1999) *Structure and Mechanism in Protein Science* (Freeman, New York).
5. Weber G (1992) *Protein Interactions* (Chapman and Hall, New York).
6. Bryngelson JD, Wolynes PG (1987) Spin glasses and the statistical mechanics of protein folding. *Proc Natl Acad Sci USA* 84:7524–7528.
7. Pande VS, Grosberg AY, Tanaka T (1997) Statistical mechanics of simple models of protein folding and design. *Biophys J* 73:3192–3210.
8. Plotkin SS, Wang J, Wolynes PG (1996) Correlated energy landscape model for finite, random heteropolymers. *Phys Rev E: Stat Phys, Plasmas, Fluids, Relat Interdiscip Top* 53:6271–6296.
9. Shakhnovich EI, Gutin AM (1993) Engineering of stable and fast-folding sequences of model proteins. *Proc Natl Acad Sci USA* 90:7195–7199.
10. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins* 21:167–195.
11. Leopold PE, Montal M, Onuchic JN (1992) Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proc Natl Acad Sci USA* 89:8721–8725.
12. Miyashita O, Onuchic JN, Wolynes PG (2003) Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. *Proc Natl Acad Sci USA* 100:12570–12575.
13. Okazaki K, et al. (2006) Multiple-basin energy landscapes for large-amplitude conformational motions of proteins: Structure-based molecular dynamics simulations. *Proc Natl Acad Sci USA* 103:11844–11849.
14. Hegler JA, Weinkam P, Wolynes PG (2008) The spectrum of biomolecular states and motions. *HFSP J* 2:307–313.
15. Wolynes PG (1996) Symmetry and the energy landscapes of biomolecules. *Proc Natl Acad Sci USA* 93:14249–14255.
16. Monod J, Wyman J, Changeux JP (1965) On the nature of allosteric transitions: A plausible model. *J Mol Biol* 12:88–118.
17. Ferreiro DU, Hegler JA, Komives EA, Wolynes PG (2007) Localizing frustration in native proteins and protein assemblies. *Proc Natl Acad Sci USA* 104:19819–19824.
18. Daily MD, Gray JJ (2007) Local motions in a benchmark of allosteric proteins. *Proteins* 67:385–399.
19. Whitford PC, Onuchic JN, Wolynes PG (2008) Energy landscape along an enzymatic reaction trajectory: Hinges or cracks? *HFSP J* 2:61–64.
20. Popovych N, Sun S, Ebright RH, Kalodimos CG (2006) Dynamically driven protein allostery. *Nat Struct Mol Biol* 13:831–838.
21. Papoian GA, et al. (2004) Water in protein structure prediction. *Proc Natl Acad Sci USA* 101:3352–3357.
22. Papoian GA, Ulander J, Wolynes PG (2003) Role of water mediated interactions in protein-protein recognition landscapes. *J Am Chem Soc* 125:9170–9178.
23. Gunasekaran K, Ma B, Nussinov R (2004) Is allostery an intrinsic property of all dynamic proteins? *Proteins* 57:433–443.
24. Henzler-Wildman KA, et al. (2007) Intrinsic motions along an enzymatic reaction trajectory. *Nature* 450:838–844.
25. Olsson U, Wolf-Watz M (2010) Overlap between folding and functional energy landscapes for adenylate kinase conformational change. *Nat Commun* 1:111.
26. Whitford PC, Miyashita O, Levy Y, Onuchic JN (2007) Conformational transitions of adenylate kinase: Switching by cracking. *J Mol Biol* 366:1661–1671.
27. Lu Q, Wang J (2008) Single molecule conformational dynamics of adenylate kinase: Energy landscape, structural correlations, and transition state ensembles. *J Am Chem Soc* 130:4772–4783.
28. Mezard M, Parisi G, Virasoro MA (1987) *Spin Glass Theory and Beyond* (World Scientific, London).
29. Li W, Wolynes PG, Takada S (2011) Frustration, specific sequence dependence, and nonlinearity in large-amplitude fluctuations of allosteric proteins. *Proc Natl Acad Sci USA* 108:3504–3509.
30. Zhang HJ, Sheng XR, Pan XM, Zhou JM (1997) Activation of adenylate kinase by denaturants is due to the increasing conformational flexibility at its active sites. *Biochem Biophys Res Commun* 238:382–386.
31. Zhuravlev PI, Papoian GA (2010) Protein functional landscapes, dynamics, allostery: A tortuous path toward a universal theoretical framework. *Q Rev Biophys* 43:295–332.
32. Kuriyan J, Eisenberg D (2007) The origin of protein interactions and allostery in colocalization. *Nature* 450:983–990.
33. Gould SJ, Lewontin RC (1979) The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proc R Soc Lond, Ser B* 205(1161):581–598.