Cell

Hot Spots for Allosteric Regulation on Protein Surfaces

Kimberly A. Reynolds,[1] Richard N. McLaughlin,[1,2] and Rama Ranganathan[1,*]
[1]Green Center for Systems Biology and Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, TX 75390-9050, USA
[2]Present address: Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
*Correspondence: rama.ranganathan@utsouthwestern.edu
DOI 10.1016/j.cell.2011.10.049

SUMMARY

Recent work indicates a general architecture for proteins in which sparse networks of physically contiguous and coevolving amino acids underlie basic aspects of structure and function. These networks, termed sectors, are spatially organized such that active sites are linked to many surface sites distributed throughout the structure. Using the metabolic enzyme dihydrofolate reductase as a model system, we show that: (1) the sector is strongly correlated to a network of residues undergoing millisecond conformational fluctuations associated with enzyme catalysis, and (2) sector-connected surface sites are statistically preferred locations for the emergence of allosteric control in vivo. Thus, sectors represent an evolutionarily conserved "wiring" mechanism that can enable perturbations at specific surface positions to rapidly initiate conformational control over protein function. These findings suggest that sectors enable the evolution of intermolecular communication and regulation.

INTRODUCTION

Allosteric regulation enables the activity of one site on a protein to modulate function at another spatially distinct site (Cui and Karplus, 2008; Luque et al., 2002; Monod et al., 1965; Smock and Gierasch, 2009). This biochemical property is fundamental to many cellular processes—in different contexts, it represents information flow between functional surfaces on signaling proteins, regulation of protein activities through molecular interactions and posttranslational modification, and functional cooperativity within oligomeric or multidomain proteins. Much prior work has examined the physical mechanism of allostery in proteins, with the finding that allostery involves the cooperative action of groups of amino acids such that local perturbations at one site can influence the function of distant sites (Clarkson et al., 2006; Luque et al., 2002). For example signal transduction in G protein-coupled receptors (Gether, 2000; Menon et al., 2001), voltage-dependent activation of K^+ channels (Lee et al., 2009; Sadovsky and Yifrach, 2007; Yifrach and MacKinnon, 2002), regulation of ligand binding in PDZ domains (Peterson et al., 2004), and modulation of catalytic rate in several enzymes (Agarwal et al., 2002; Benkovic and Hammes-Schiffer, 2003; Eisenmesser et al., 2005; Fraser et al., 2009) all seem to depend on networks of functionally coupled residues that exist within the overall atomic structure. The mechanistic details vary, but the salient point is that allostery involves cooperative interactions between a subset of spatially distributed amino acids.

One approach for understanding the structural basis of functional properties like allostery is the analysis of amino acid coevolution using methods such as statistical coupling analysis (SCA) (Halabi et al., 2009; Lockless and Ranganathan, 1999; Süel et al., 2003). The basic premise of SCA is that functionally relevant coupling between amino acids should, regardless of underlying mechanism, drive coevolution of those residues. In principle the global pattern of coupling between amino acids can be estimated from multiple sequence alignments that represent the long-term evolutionary record of a protein family.

This approach reveals two general findings about proteins. First, the majority of amino acids in proteins evolve nearly independently, implying weak or idiosyncratic physical coupling to other residues. This result is nontrivial; many of these weak interactions constitute direct contacts in the tertiary structure, suggesting extensive decoupling even within local environments. Second, a small fraction of amino acids (typically 10%–30%) show strong mutual coevolution and comprise spatially distributed but structurally contiguous subnetworks within the tertiary structure. These coevolving networks are termed "sectors," and experiments in several protein families confirm that sectors are associated with conserved functional properties, including signal transmission, allosteric regulation, and catalysis (Ferguson et al., 2007; Halabi et al., 2009; Hatley et al., 2003; Shulman et al., 2004; Süel et al., 2003). In addition for a small protein interaction module, computational design of synthetic proteins based on the pattern of evolutionary couplings between amino acids was shown sufficient to recapitulate the structure and function of their natural counterparts (Russ et al., 2005; Socolich et al., 2005). Multiple sectors are possible in a single protein (Halabi et al., 2009), arguing that this architecture permits the independent variation of different phenotypes that comprise overall fitness. Thus, sectors are proposed to represent the fundamental structural units that underlie the conserved structure and functions of natural proteins.

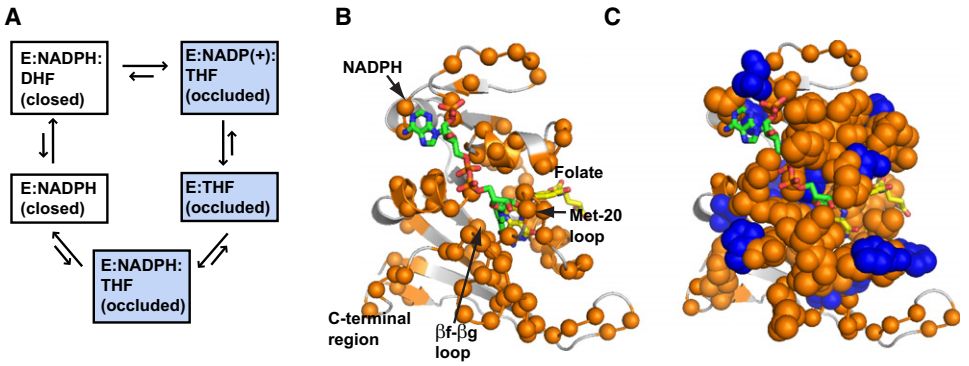


Figure 1. The DHFR Sector and Residues Involved in Millisecond Dynamics Relevant to Catalysis

(A) The reaction cycle of *E. coli* DHFR. DHFR catalyzes the stereospecific reduction of DHF to THF through transfer of a hydride ion from the cofactor NADPH. The main structure change associated with the reaction cycle is a switch between the so-called closed and occluded conformations, a fluctuation that occurs on a similar timescale as the catalytic step of the reaction.

(B) A mapping of amino acids undergoing conformational exchange at the millisecond timescale in any of the complexes of *E. coli* DHFR representing the catalytic cycle (shown as small orange spheres on the Cα atom; PDB 1RX2) (Boehr et al., 2006, 2010; McElheny et al., 2005).

(C) The SCA sector for the DHFR family is shown in CPK (blue and orange; see Table S1). Orange and blue spheres represent sector positions either overlapping or not, respectively, with residues undergoing millisecond dynamics. The small orange spheres represent nonsector positions involved in millisecond dynamics. This analysis shows that sector positions are strongly correlated with residues undergoing dynamic motions underlying catalysis ($p < 0.006$; see text and Table S2).

In addition to folding and function, natural proteins display the capacity for evolving novel allosteric regulation and communication (Kuriyan and Eisenberg, 2007). The origin of such regulatory mechanisms is not obvious given the complexity of building cooperative interactions between amino acids that can functionally couple distant sites. Interestingly, the sector architecture suggests a simple potential solution. Sectors have a distributed spatial organization that "wires" the active site to multiple distant surface positions. This architecture provides constraints for structure and function (Russ et al., 2005; Socolich et al., 2005) but, as a consequence, might also provide a structural basis for the gain of novel allosteric regulation through initiation of molecular interactions at sector-connected surface sites. In other words, sector-connected surface sites might represent "hot spots" for the emergence of allosteric control in proteins.

Here, we experimentally test this hypothesis using the metabolic enzyme dihydrofolate reductase (DHFR) and a protein interaction module (the PDZ domain) as model systems. We show that DHFR contains a sparse, distributed, and physically contiguous protein sector that is strongly correlated to the dynamic motions associated with catalysis. By carrying out a comprehensive domain insertion scan in *E. coli* DHFR and applying a new assay system, we show that sector-connected surface sites are indeed hot spots for the emergence of allosteric control. We recapitulate this finding in a second experimental system, the PDZ domain. Interestingly, initiation of molecular interactions at these sector-connected sites can produce allosteric regulation in a single step that is detectable in vivo, without directed optimization or design. These results show that sectors can provide a statistically preferable route for the initiation of allosteric control and suggest that they can enable the evolution of allosteric communication between proteins.

RESULTS

A Link between the Sector and Functional Conformational Dynamics in DHFR

DHFR is an essential enzyme in both prokaryotes and eukaryotes that is necessary for the biosynthesis of purines, pyrimidines, and amino acids. The enzyme catalyzes the stereospecific reduction of 7,8-dihydrofolate (DHF) to 5,6,7,8-tetrahydrofolate (THF) using nicotinamide adenine dinucleotide phosphate (NADPH) as a cofactor and has served as an excellent system for understanding catalysis and the relationship between conformational dynamics and enzyme activity (Schnell et al., 2004). Specifically, the *E. coli* DHFR reaction cycle involves five biochemically characterized catalytic intermediates that comprise two major conformational states, termed closed and occluded (Figure 1A) (Fierke et al., 1987; Schnell et al., 2004). NMR-based relaxation dispersion experiments show that a distributed pattern of motion on a millisecond timescale plays a crucial role in mediating the conformational transitions between the various states that mediate substrate and cofactor binding/release and the chemical step in catalysis (Bhabha et al., 2011; Boehr et al., 2006, 2010; McElheny et al., 2005). Considered collectively, the set of DHFR residues engaged in millisecond scale dynamics encompasses the active site, substrate and cofactor binding sites, and some distant regions (Figure 1B). Thus, the catalytic mechanism in DHFR involves millisecond dynamics within a distributed network of amino acid residues.

SCA for an alignment of 418 phylogenetically diverse DHFR sequences defines a system of coevolving amino acid positions (a sector) that depending on statistical cutoffs, comprises between 14% and 31% of the protein (see Table S1 available online and Experimental Procedures). Consistent with prior descriptions, the sector forms a physically contiguous network of atoms that connects the DHFR active site with the substrate

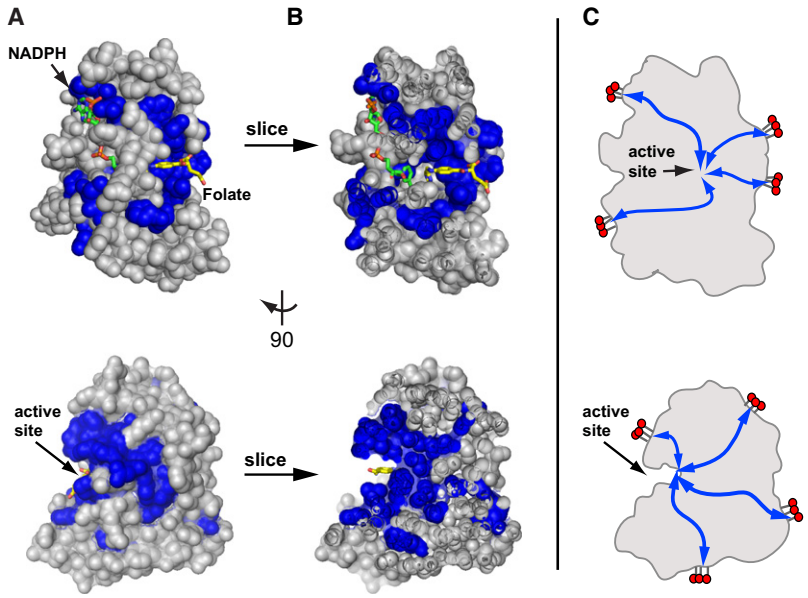


Figure 2. Sector Architecture in DHFR

(A and B) Two views (different by 90° rotation) of the surface (A) and a slice through the protein core (B) with sector residues colored in blue. Substrate and cofactor are shown in yellow and green stick bonds, respectively.

(C) A cartoon representation of the slice mappings in (B), illustrating that the sector comprises a sparse, physically connected network of residues that link the active site to a few distant surface positions (red). See also Table S1.

and cofactor binding pockets, and with several distantly positioned surface regions (Figures 1C and 2) (Chen et al., 2007; Lee et al., 2008). Thus, despite no known allosteric function, DHFR contains the same sparse and distributed sector architecture as other proteins.

Comparison of the SCA sector with those residues undergoing millisecond timescale dynamics in DHFR shows a strong coincidence of the two; nearly 75% of sector positions overlap with those showing motions related to the catalytic cycle (Figure 1C; and for a full list of sector positions, see Table S1). This corresponds to a strong statistical correlation between sector positions and those residues involved in millisecond dynamics ($p < 0.006$ by Fisher's exact test), a result that is robust to perturbations in the cutoffs used for sector identification (Table S2). These data suggest that (1) the sector in DHFR represents an evolutionary-conserved distributed architecture involved in the catalytic reaction cycle; and (2) the sector mechanistically operates through dynamic fluctuations that connect the active site with several distant surface sites.

A Comprehensive Test of Allosteric Regulation at Surface Positions

What does the sector architecture in DHFR mean for the capacity of surface positions to initiate functional control over catalytic function? The finding that the sector connects the active site to a number of surface positions suggests the idea that by virtue of distributed constraints on catalysis, the sector provides a preorganized path for coupling between these distantly positioned sites (Figure 2). If true, initiation of new molecular interactions at sector-connected surface sites should preferentially trigger the emergence of allosteric regulation. The initial magnitude of regulation might be weak given no optimization or intelligent engineering, but to be evolutionarily significant, need only be sufficient to provide a basis for selection in vivo.

To examine this, we carried out a test for the emergence of allosteric control at all surface-exposed sites in *E. coli* DHFR (Figure 3A) using a technique we term "domain insertion scanning." The strategy is to create a library of chimeric DHFRs in which an unrelated allosteric signaling module (a light-sensitive PAS domain [LOV2] from *A. sativa*; Salomon et al., 2001) is inserted into the peptide bond preceding every solvent-exposed residue in DHFR (70 total, Figure 3A; see Experimental Procedures). The N- and C-terminal regions of the LOV2 domain are known to receive an allosteric conformational change upon photon absorption by the bound flavin mononucleotide chromophore (Halavaty and Moffat, 2007; Harper et al., 2003). More specifically, a C-terminal helix (referred to as the Jα helix) detaches from the core of the LOV2 domain in response to light. Thus, inserting the LOV2 domain at each DHFR surface site and assaying for the emergence of light-dependent DHFR activity represent a test for the initiation of allosteric regulation. A previous proof-of-principle experiment establishes this experimental approach; insertion of LOV2 at one sector-connected surface site (120–121, here called DL121) results in weakly light-dependent DHFR activity in vitro (Lee et al., 2008).

To assay light-dependent DHFR activity of all 70 DHFR-LOV2 chimeras in vivo and in parallel, we developed a folate auxotroph rescue assay coupled with measurement of growth rates by Solexa-based high-throughput sequencing. THF, the product of DHFR catalysis, is needed for a number of critical metabolic processes, including the synthesis of thymidine and amino acids. As a result, DHFR catalysis is necessary for growth of *E. coli* (Rajagopalan and Benkovic, 2002). For example the *E. coli* folate auxotroph (ER2566 *ΔfolA ΔthyA*, which contains deletions of both DHFR [*folA*] and thymidylate synthetase [*thyA*]) fails to grow in minimal media conditions but can be rescued with basal expression of both the *folA* and *thyA* genes from a plasmid. To test if the insertion of LOV2 abrogated DHFR catalytic activity for any of the fusions, we screened the library of DHFR-LOV2 chimeras for those that provide measurable rescue of auxotroph growth. We find that 67 of the 70 chimeric fusions can complement, indicating the presence of a functional DHFR (Figure S1). The three insertions lacking DHFR activity (DL17, DL23, and DL108) were omitted from further analysis.

A study of a few known mutants of DHFR shows that the growth rate is a monotonic function of the catalytic efficiency (k_{cat}/K_m) in the conditions of our experiment (Figure S2). How

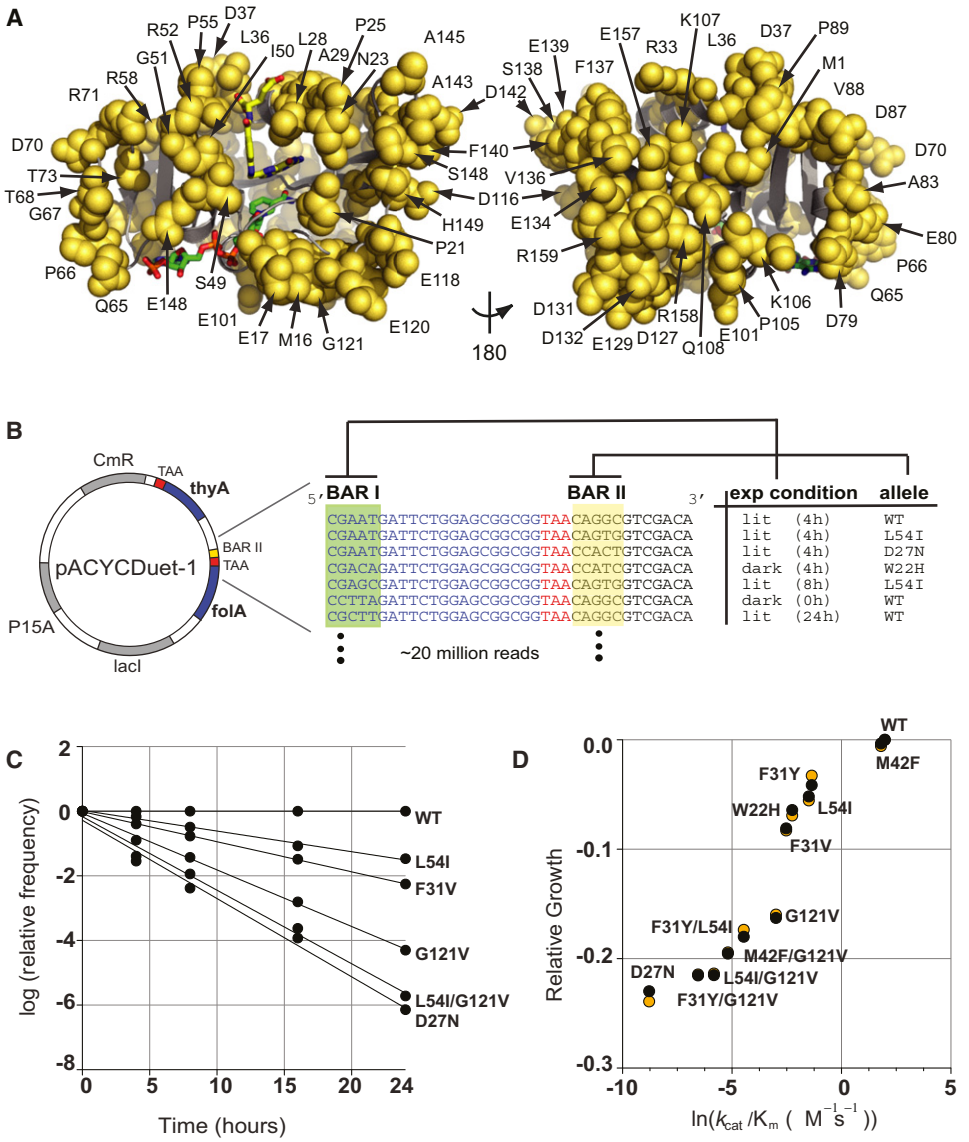


Figure 3. Comprehensive Domain Insertion Scan and Relative Growth Rate Measurements by High-Throughput Sequencing

(A) All LOV2 domain insertion sites on the DHFR surface (70 in total, orange spheres). For simplicity in discussion we refer to each DHFR-LOV2 chimera or DHFR mutant as a "variant."

(B) Barcoding strategy for the DHFR variants. Each DHFR mutant or DHFR-LOV2 chimera was labeled with two DNA barcodes: (1) a 5 bp barcode that identifies the time point of sampling and experimental (exp) condition (dark or lit); and (2) a 5 bp barcode immediately following the DHFR stop codon that identifies the variant. The first barcode was added to the 5′ end of the sequenced region during sample preparation by PCR (see Experimental Procedures). Sequencing of both barcodes permits determination of relative variant frequencies within a mixed population as they vary with time and experimental condition.

(C) Measurement of growth rates through sequencing for a set of DHFR point mutants that span a broad range of catalytic activities in vitro. The log frequency of each variant is shown relative to WT, and is normalized to the initial values at the start of the experiment (t = 0; see Experimental Procedures). Thus, slopes of the linear regression report growth rates relative to WT.

(D) Comparison between in vitro catalytic power and in vivo relative growth rate, indicating a monotonic relationship between the two (see Table S3; and Figures S1 and S2). Yellow and black circles represent two independent experimental trials in the light and dark, respectively.

can we make accurate and systematically controlled measurements of light-dependent growth rate for all chimeras with reasonable experimental speed? Previous work has demonstrated that even small differences in growth rates of two variants can be measured through pairwise competition of fluorescently labeled strains (Breslow et al., 2008; Thompson et al., 2006). Here, we introduce a new method for monitoring relative fitness by measuring individual variant frequencies over time in a mixed population of growing cells using the technology of massively parallel sequencing (Metzker, 2010). Briefly, ER2566 *ΔfolA ΔthyA* cells expressing all 70 DHFR-LOV2 chimeras, as well as the wild-type (WT) DHFR, a set of 11 DHFR mutants, and 2

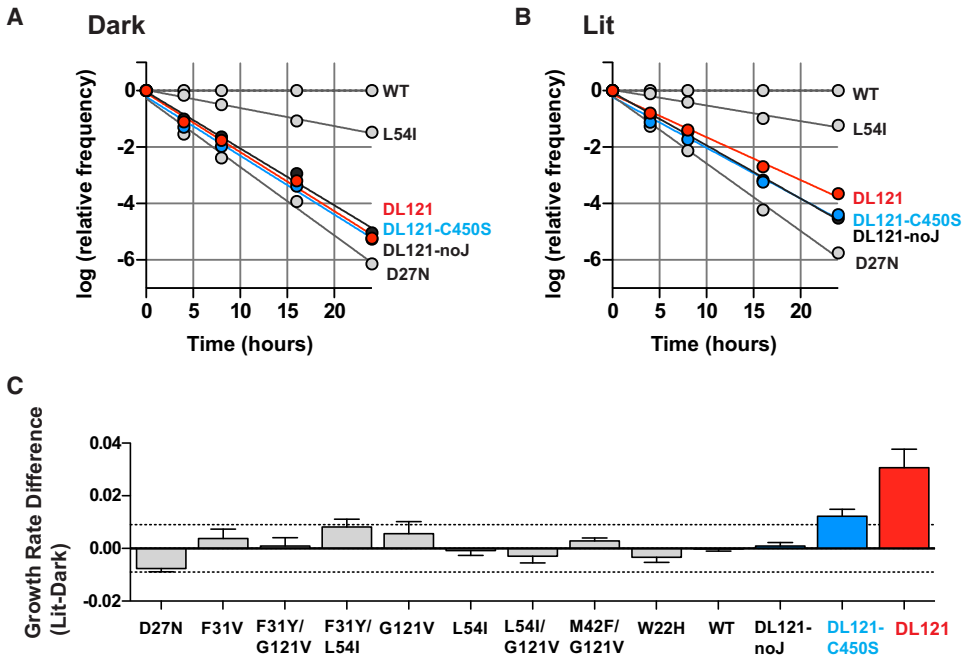


Figure 4. Light Dependence in Growth Rate for One Chimera, DL121, Previously Shown to Display Weak Light Dependence in Catalytic Rate In Vitro

(A and B) Experiments under dark and lit conditions show that DL121 displays an ~16% increase in growth rate in response to light, whereas two DL121 variants carrying LOV2 domains defective in allosteric mechanism (121-C450S and 121-noJ) do not show light dependence (Figure S3).

(C) Quantitative measurement of light dependence in growth rate in three independent growth/sequencing experiments for ten nonlight-dependent DHFR mutants spanning a broad range of catalytic power, and for DL121, 121-C450S, and 121-noJ. Error bars indicate the SEM across the three experiments. The data demonstrate good reproducibility in growth rate measurements in independent experiments, and establish a statistical model for measurement noise in this assay based on the behavior of the nonlight-dependent mutants of DHFR (the dashed lines indicate 2σ from the mean). In comparison, DL121 shows clear statistically significant light dependence, whereas 121-C450S and 121-noJ do not.

additional DHFR-LOV2 control constructs (described in detail later) were mixed in equivalent ratios (84 variants total), grown in a single flask under either dark or lit conditions, and sampled at 5 time points. To identify the individual variants, we incorporated a five-nucleotide barcode at the 3′ untranslated region downstream of the *folA* gene (Figure 3B). In addition during sample preparation for Solexa sequencing, we added a second five-nucleotide barcode encoding the experimental condition (lit or dark, and each time point) to the 5′ end of the sequenced region by PCR (Figure 3B). Multiplexing all samples and sequencing a 36 bp fragment containing both barcodes permit reconstruction of the growth rate divergence of every variant over time with excellent counting statistics (Figure 3C). This experiment enables simultaneous growth rate determinations for a very large number of variants in a single internally controlled experiment without the need for external labeling of individual strains (Figure 3D).

To establish the growth-based sequencing assay as a quantitative reporter of DHFR activity, we expressed and purified the 11 DHFR mutants to near homogeneity and measured catalytic power in vitro using standard protocols (Figure S3 and Table S3) (Rajagopalan et al., 2002). Consistent with the conventional growth rate measurements by optical density (Figure S2), there is a monotonic and sensitive relationship between DHFR catalytic activity in vitro and growth rate of *E. coli* (Figure 3D). Indeed, DHFR mutants L54I and W22H differ only 2-fold in in vitro activity (as assessed by k_{cat}/K_m), but this results in an approximately 12% growth rate advantage for W22H over L54I. As previously shown, the exponential divergence of populations with differences in growth rate makes it so that even small biochemical effects in vitro can be accurately detected in vivo (Breslow et al., 2008). The measurements are also highly reproducible; over the large range of catalytic activities sampled by the ten nonlight-dependent DHFR mutants, there is negligible difference in the lit and dark growth rates measured in three independent growth/sequencing experiments. The distribution of differential growth in lit and dark conditions is centered at zero, with a small variance that we attribute to measurement noise (0.0065 ± 0.0047, mean ± SD [σ]; Figures 3D, 4C, 5A, and 5B).

To establish the degree to which we expect light-dependent DHFR catalytic activity in vitro to control growth rate in response to light in vivo, we considered one previously characterized DHFR-LOV2 chimera (DL121) (Lee et al., 2008). In biochemical experiments, DL121 displays a modest 2-fold light dependence in k_{hyd} (the rate of hydride transfer, the chemical step of DHFR catalysis). As controls, light dependence in growth rate was also assessed for two light-insensitive variants of DL121: a point mutant in the LOV2 domain that does not undergo conformational change in response to light

(DL121-C450S); and a variant lacking the output mechanism of LOV2 (DL121-noJ, a deletion of the LOV2 Jα helix). The data show that DL121 shows a significant increase in growth rate of *E. coli* in response to light but that the neither control construct shows light dependence above measurement noise (Figure 4). This effect is reproducible and robust; multiple trials of the sequencing-based experiment and an independent assay (fluorescence-based measurement of relative growth rates) show similar results (Figure S4). The overall effect for bacteria carrying DL121 is an ~17% increase in growth rate in the light compared to the dark, a finding that shows how subtle biochemical allostery in vitro can translate to a nontrivial fitness advantage in vivo.

Allosteric Regulation Occurs Preferentially at Sector-Connected Surfaces

We measured lit and dark growth rates for all 70 of the DHFR-LOV2 fusions in the 3 independent growth/sequencing experiments. A few chimeras grew unreliably in the experiment trials and were removed from consideration, leaving 61 chimeras and 10 light-independent controls for further analysis (Figure S5). Light dependence was quantitatively assessed by calculating a standard Z score that indicates the deviation of the lit minus dark growth rate from that measured for the light-independent controls (see Experimental Procedures). Like the DHFR point mutants, most DHFR-LOV2 chimeras show light dependence close to zero (Figures 5A–5C). This indicates that the majority of surface sites do not show allosteric regulation upon insertion of LOV2. Purification and in vitro spectral analysis of a sampling of nonlight-dependent DHFR-LOV2 chimeras indicate an active, chromophore bound LOV2 domain displaying light-dependent dynamics similar to that of the WT domain (Figure S6). Together with evidence for intact DHFR activity (Figure S1), these data support the conclusion that lack of light dependence is due to the functional uncoupling of DHFR and LOV2 rather than defects intrinsic to either domain.

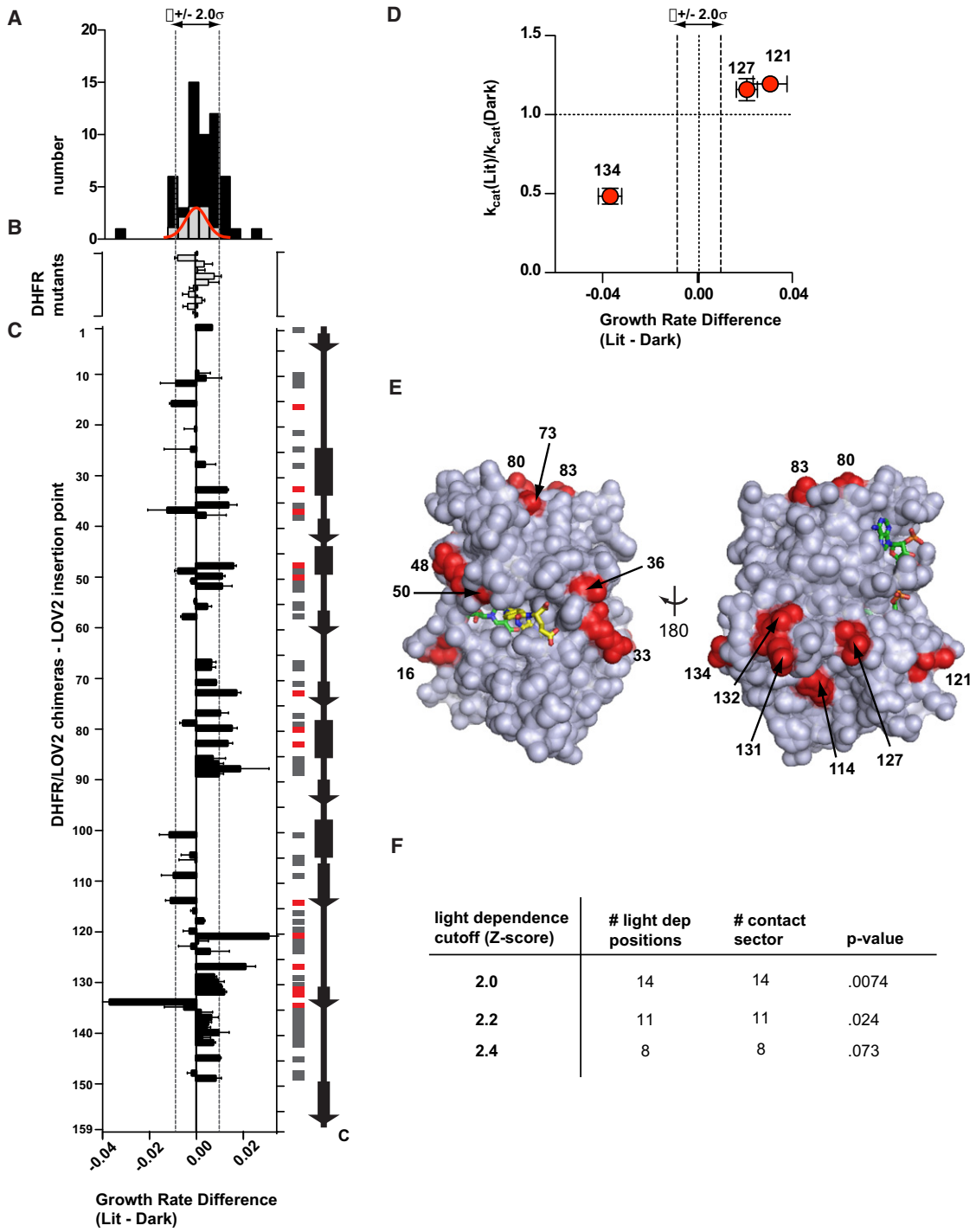
However, the distribution of light-mediated differences in growth rate (Figure 5A) indicates a light dependence ($Z > 2$) over nonlight-dependent controls for 14 out of 61 DHFR insertion sites (Figures 5A–5C and Table S4). These chimeras display growth rate differences between the lit and dark conditions that range from a few to tens of percent, suggesting that like for DL121, LOV2 insertion triggers weak but significant allosteric control in vivo at a subset of surface positions. For a sampling of DHFR-LOV2 fusions that span the full range of light dependence in growth rate, we purified the chimeric proteins to near homogeneity and measured the catalytic rate of DHFR (k_{cat}) under lit and dark conditions (Figure 5D). These measurements demonstrate that the light dependence of growth rate in vivo is highly correlated to the light dependence of enzymatic activity in vitro, a result that argues that insertion of LOV2 acts directly to modulate DHFR activity.

The sites showing significant allosteric regulation are distributed throughout the primary and secondary structure of the protein (Figure 5C), and occur at surface locations with no obvious spatial relationship to the active site or to each other (Figures 5E and 6, red spheres). Indeed, proximity to the active site is a poor predictor of light-dependent regulation; 5 out of 14 light-dependent positions occur within 10 Å of the active site, from a total of 28 LOV2 insertion sites within this distance ($p < 0.28$, Fisher's exact test). However, we find that every one of the light-dependent surface sites is connected to the DHFR sector (Figure 6, blue spheres), a result that indicates strong correlation between light dependence and sector connectivity ($p < 0.007$, Fisher's exact test with 2σ cutoff for light dependence). The statistical significance of sector connectivity holds over a range of the tail of the light-dependent distribution (Figure 5F and Table S5), especially including more stringent definitions of light dependence that are less susceptible to measurement noise. In addition sector connectivity of light-dependent positions also holds for a broad range of cutoffs used for sector definition (Table S5). Indeed, light-dependent positions are even more significantly associated with sector positions with the strongest correlation signals, a finding that provides further confidence in the association of the two. Thus, the data strongly support the proposal that sector-connected surface sites are hot spots for the initiation of allosteric regulation.

The Spatial Architecture of Allosteric Control

Slices through the core of DHFR illustrate the structural organization of sector positions and LOV2 insertion sites showing light dependence (Figure 6). The sector is like a preorganized functional "wire" in the three-dimensional structure that connects every light-dependent position to the enzyme active site through pathways of intervening residues. As previously reported, the sector is not any obvious property of primary, secondary, or even tertiary structure; it is a network of mutually evolving residues that presumably emerges from the heterogeneity of cooperative interactions between amino acids (Halabi et al., 2009). The sector in DHFR represents the distributed determinants of catalytic mechanism (Figure 1C), but as a consequence, also provides the capacity for gaining regulation at neighboring surface positions through single-step variation (in this case, domain insertion). The spatial pattern of sectors enables multiple surface positions to participate, a finding that is consistent with the observed diversity of allosteric mechanisms within a protein family (Kuriyan and Eisenberg, 2007).

The association of the sector with residues engaged in functionally relevant dynamic motions in DHFR suggests a possible mechanistic basis for light-dependent allosteric control. Out of the 14 light-dependent positions, 13 contact the network of residues undergoing millisecond conformational exchange (the 1 noncontacting residue is indirectly connected to this network via another light-dependent position) (data not shown). Thus, a working model is that sector-mediated allosteric regulation in DHFR works through modulation of dynamic modes associated with catalysis. This observation provides a starting point for understanding the physical basis for allostery through sector connectivity, but we emphasize that the mechanistic details of allosteric control are likely to vary from protein to protein and, indeed, from site to site. Thus, sector connectivity is a phenomenological principle for the emergence of allostery that we expect will be implemented through a variety of degenerate physical mechanisms.



F

light dependence cutoff (Z-score)	# light dep positions	# contact sector	p-value
2.0	14	14	.0074
2.2	11	11	.024
2.4	8	8	.073

Figure 5. The Emergence of Allosteric Control at Sector-Connected DHFR Surface Sites

(A) Histograms of growth rate difference (lit-dark) for all DHFR-LOV2 chimeras (black) and nonlight-dependent DHFR mutants (gray, with Gaussian fit in red). (B and C) Growth rate differences for nonlight-dependent mutants (B) and for DHFR-LOV2 chimeras ordered by DHFR primary structure (C) (see also Figure S4). Error bars indicate SEM across three experimental repeats. The corresponding secondary structure pattern is indicated at right. The gray/red bars at right indicate the position of each LOV2 insertion; red bars indicate insertions sites showing light dependence ($Z > 2$) of nonlight-dependent controls (Table S4). Light-dependent positions are scattered throughout the primary and secondary structure of the protein.

(D) Correlation of light dependence in vitro and in vivo. The catalytic rate (k_{cat}) was measured for four DHFR-LOV2 fusions under lit and dark conditions; error bars represent the standard deviation (σ) across three experimental trials. This confirms that DL121, DL127, and DL134 show light dependence in enzymatic activity as measured biochemically.

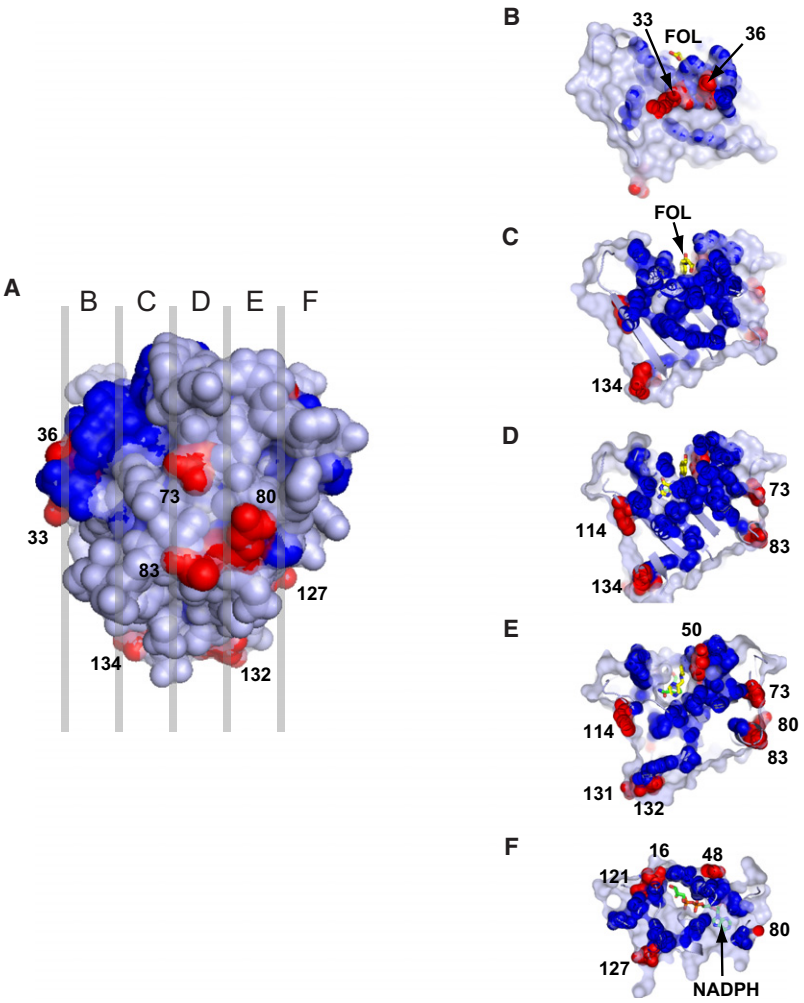


Figure 6. Pathways of Sector Connectivity between the Active Site and Light-Dependent Surface Positions

(A) Space-filling representation of DHFR with light-dependent surface positions in red.

(B–F) Serial slices taken through DHFR at the planes indicated in (A); the views in (B)–(F) are from the left. The data show that sectors form physically contiguous pathways through the core of the three-dimensional structure that connect all light-dependent positions with the substrate and cofactor binding sites and with the catalytic active site. Substrate and cofactor are shown as yellow or green stick bonds, respectively. Thus, light-dependent positions are "wired up" to the active site through sector amino acids.

Sector-Mediated Hot Spots for Allosteric Regulation in the PDZ Domain

Is the concept of allosteric control through sector connectivity a general phenomenon in proteins? To begin testing this, we carried out the same experiment—a comprehensive surface scan for positions capable of displaying functional regulation—in a different model system, the PDZ family of protein interaction modules. PDZ domains are roughly 100 amino acid proteins that typically bind the carboxy-terminal few amino acids of various target proteins and are components of multidomain scaffolding complexes (Nourry et al., 2003). Recent technical advances make it possible to quantitatively evaluate the functional impact of mutating every position in the PDZ domain to every other amino acid (R.N.M. and R.R., unpublished data). Carried out for all surface-exposed amino acid positions, this analysis provides the opportunity to reexamine the hypothesis that allosteric regulation of protein function is mediated through the distributed connectivity of sector amino acids.

We mutated every surface-exposed position in 1 specific PDZ domain ($PSD95^{pdz3}$, 39 total positions) to every other amino acid, and measured the effect on binding its cognate ligand, the C-terminal peptide from the CRIPT protein (Niethammer et al., 1998; Figure S7). The data show that 11 of the 39 surface positions have significant effects on ligand binding upon mutation (Figure S7), and nearly every one (10 of 11) is sector connected (Figure 7; $p < 0.039$, Fisher's exact test). Indeed, the data essentially recapitulate the result from LOV2 insertion in DHFR; functionally coupled surface positions are distributed at surfaces of the PDZ domain with no obvious spatial relationship to the binding site or to each other but are linked via pathways of sector residues through the protein core (Figure 7).

Taken together, the data in DHFR and PDZ suggest the model that, by preorganization of a few amino acid positions into collectively acting units, the sector represents an architecture that is capable of initiating allosteric control in proteins.

DISCUSSION

Protein Sectors and Capacity for Novel Allosteric Regulation

Despite fundamental importance in nearly every biological process, the general structural principles behind the mechanisms and evolutionary origin of allosteric communication in proteins have been difficult to elucidate. The basic problem

(E) Mapped on an atomic structure of *E. coli* DHFR, light -dependent positions (red) comprise a spatially distributed subset of the protein surface positions (light blue).

(F) Statistics of DHFR positions showing sector connectivity and light dependence. Every light-dependent (dep) position is also sector connected over a range of significance thresholds for light dependence and sector definition (Table S5). These results indicate robust statistical correlation between sector connectivity and capacity for allosteric control.

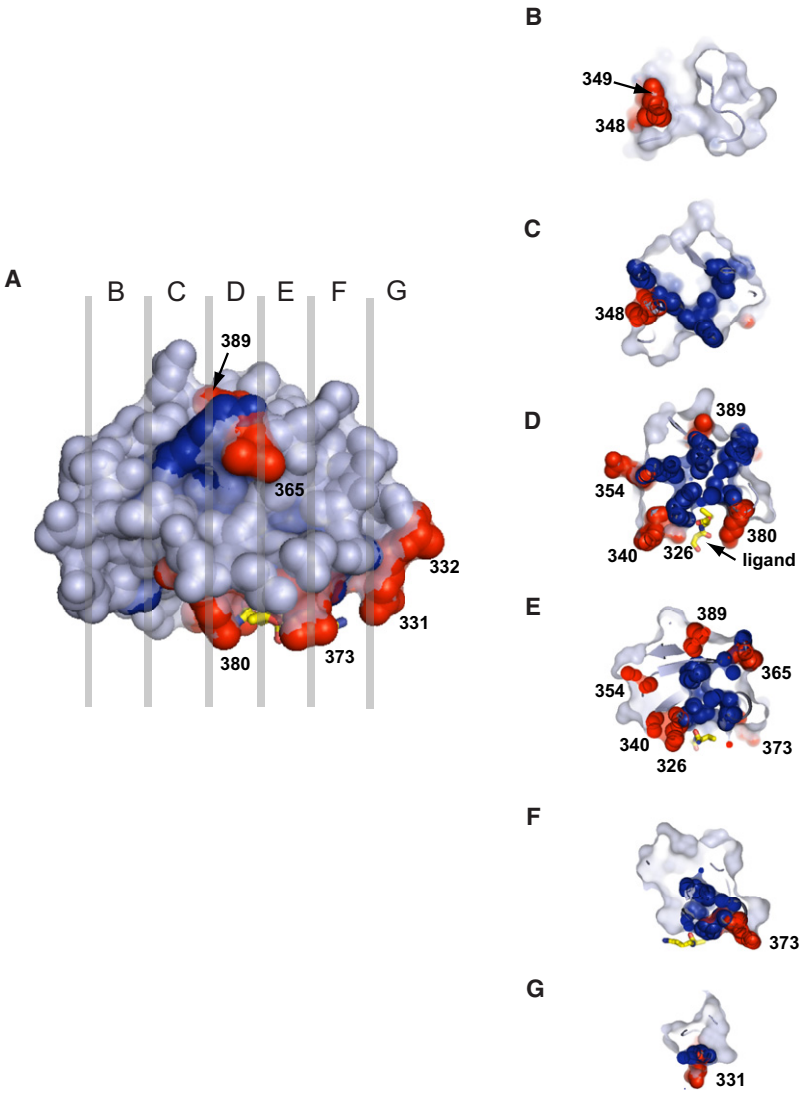


Figure 7. Sector-Connected Surfaces Modulate Function in the PDZ Domain

(A) Space-filling representation of PDZ with surface mutations that perturb protein function in red.

(B–G) Serial slices taken through PDZ at the planes indicated in (A); the views in (B)–(G) are from the left. As for DHFR, the data show that sectors form physically contiguous pathways through the structure that connect all mutations that impact PDZ function to the peptide binding site. The peptide is indicated in yellow stick bonds.

has been the difficulty of inferring the net functional value of interactions between amino acids from the pattern of observed contacts in protein structures. Taking advantage of the sequence divergence in protein families to formulate a statistical approach to this problem, we previously proposed the model that natural proteins have a general "design" in which sparse, distributed, and physically contiguous networks of strongly coupled amino acids (sectors) are embedded in an overall environment of weak coupling (Halabi et al., 2009; Lockless and Ranganathan, 1999). Experiments in several model systems argue that sectors are linked to protein function, including long-range communication between protein surfaces. Here, we show that the metabolic enzyme DHFR, a protein with no known allosteric or signaling function, has a protein sector that is architecturally no different: a subset of total amino acids comprise a contiguous network of residues in the protein core that links the active site to a number of distantly positioned surface sites. Functionally, the DHFR sector corresponds well to residues involved in the catalytic mechanism, and mechanistically, it correlates with positions that undergo conformational dynamics associated with the reaction coordinate of the enzyme. These findings represent, to our knowledge, the first instance of a physical mechanism underlying the distributed connectivity that characterizes the sector architecture.

Quantitative comparison of sector edges and light-dependent insertion sites confirms that sector-connected surfaces are statistical hot spots for the initiation of allosteric control. It is important to underscore that this conclusion is probabilistic; although every light-dependent site is sector connected, not every sector-connected surface site shows allosteric control in our experiment. This is expected given that domain insertion was carried out naively, without physics-based design or directed evolution. Indeed, the naive coupling efficiency between the LOV2 domain and DHFR at sector-connected surface sites can be estimated from our experiment to be about 0.45. However, light dependence was never observed at nonsector-connected sites, a result that highlights the importance of sector connection in initiation of allosteric control. As with any experiment, it is possible that some LOV2 insertion sites that are statistically insignificant could actually be very weakly light dependent. But regardless, the data show that sector-connected surface sites display the greatest likelihood and magnitude of allosteric control and, thus, are statistically preferred sites for regulation.

Nevertheless, we note that even at sector-connected sites, the magnitude of allosteric control upon LOV2 insertion is invariably and expectedly weak. Biochemical studies show that DL121, a chimera with one of the largest light-dependent effects, only displays a roughly 2-fold increase in the microscopic catalytic rate constant (k_{hyd}) and a 20% increase in the steady-state catalytic rate (k_{cat}) in vitro upon light exposure (Lee et al., 2008). But this corresponds to a growth rate advantage in multiple different experiments in vivo of nearly 20% in the light versus dark, a value that can easily drive selection if allosteric control is a condition of fitness. In this regard we note that under the experimental conditions used in this work, the DHFR-LOV2 fusions display slower growth rates than the WT enzyme

(Figure S5) and would, therefore, be evolutionarily disadvantaged based on absolute growth rate. However, initial experiments suggest that the relationship of DHFR catalytic activity to growth rate can depend on the expression level of DHFR (K.A.R. and R.R., unpublished data), a property that is known to fluctuate naturally during the cell cycle (Almasan et al., 1995). It will be interesting to demonstrate experimental conditions in which the DHFR-LOV2 chimeras can be selected in vivo as a foundation for ultimately understanding the evolutionary processes that can select and optimize the initial allosteric effects described here.

Experimental study shows that the concept of allosteric regulation through sector connectivity is recapitulated in a second model system—the PDZ domain. This result strongly argues that the result that sector edges are allosteric hot spots does not depend on either the choice of the model system or on the manner of surface site perturbation (domain insertion or point mutagenesis). In summary the connection between sectors and initiation of allosteric control suggests that rather than emerging idiosyncratically in proteins, regulation can emerge at specific surface sites by taking advantage of proximity to preorganized cooperative networks associated with function.

A Model for the Evolution of Regulation

Empirical evidence suggests that allostery emerges readily in the evolution of proteins, often resulting in a diversity of regulatory mechanisms in members of a single protein family (Kuriyan and Eisenberg, 2007). An important step in understanding the origin of protein regulation is a theory that can explain how an evolutionary process can produce allosteric coupling with mechanistic diversity despite the complexity of constructing cooperative interactions between amino acids connecting distant functional sites. Kuriyan and Eisenberg provide one critical part of such a theory by arguing that colocalization of proteins (through recombination or compartmentalization) provides sufficient local concentration such that even single mutations at surface sites can initiate binding between proteins (Kuriyan and Eisenberg, 2007). But a key second part of the problem is to explain how interaction at random surface positions could generate functional coupling between the active sites of two proteins.

The data presented here provide a potential solution to this problem. We propose that sectors are inherent to the structure of natural proteins because they provide the basic rules for native folding and function. For example in DHFR the sector corresponds to the constraints on catalytic mechanism. Given sectors, the principle of sector connectivity suggests how recombination or compartmentalization (Kuriyan and Eisenberg, 2007) can lead to novel allosteric regulation in a single step of variation. Colocalized but independently acting protein domains can initiate the formation of novel interfaces at surface sites through single mutation. However, if allosteric coupling between the two proteins is a condition of fitness (e.g., the light dependence of DHFR), then the data presented here argue that surfaces in contact with sector residues represent statistical hot spots for interface formation. Finally, applying the Rosetta Stone principle of Marcotte and Eisenberg (Marcotte et al., 1999), further optimization of binding affinity at the newly formed interface should ultimately permit separation of protein domains through gene fission, resulting in the creation of proteins now displaying allosteric communication in *trans* through functional linkage of sectors.

Although much further work will be necessary to test this idea, evidence from natural systems supports the idea of allosteric communication through sector connectivity. The molecular chaperone Hsp70 is comprised of two allosterically coupled domains. SCA for the family of Hsp70-like proteins indicates a single sector coevolving between these two domains that physically connects the functional sites through the interdomain interface (Smock et al., 2010). In addition the regulation of binding affinity for ligands in the Par6 PDZ domain occurs through binding of the small GTPase Cdc42 at a distant allosteric site. SCA for both the G protein and PDZ families reveals a physically contiguous path of sector residues that connects the nucleotide-binding pocket in Cdc42 to the ligand-binding site in the Par6 PDZ domain through the protein-protein interface (Lee et al., 2009; Peterson et al., 2004). These examples set the stage for a more comprehensive test of the principle of sector connectivity as the basis for allosteric communication between proteins.

The model described here focuses on a plausible evolutionary path for protein-protein interactions, but this process may be generalized to the evolution of regulation by other mechanisms, such as posttranslational modification or small molecule ligand binding. In general we suggest that the sector architecture enables the rapid evolution of allosteric regulation. This model provides a foundation to understand the elaboration of complex cellular signaling and metabolic systems through systematic variation and selection.

EXPERIMENTAL PROCEDURES

Statistical Coupling Analysis

A multiple sequence alignment consisting of 418 DHFR sequences was assembled as described in Lee et al. (2008). Statistical coupling analysis (SCA) was performed as in Halabi et al. (2009), but using an updated version of MATLAB SCA toolbox (SCA Toolbox 4.0). The SCA codes and a script for computing the correlation matrix and definition of sectors through spectral decomposition are available on our laboratory web site (http://systems.swmed.edu/rr_lab/). Further details of the method are provided in the Extended Experimental Procedures.

Chimera Construction and Barcoding

The DHFR-LOV2 fusions were constructed by standard PCR stitching methods performed in two consecutive rounds of PCR with overlapping oligonucleotides. A more detailed description of the cloning strategy and choice of sequencing barcodes is provided in the Extended Experimental Procedures.

Auxotroph Rescue Assay

All relative growth rate measurements were performed in the *E. coli* folate auxotroph strain ER2566 *ΔfolA ΔthyA* (Lee et al., 2008). Selection was performed in minimal media A (described previously in Saraf et al., 2004, but without thymidine because thymidylate synthase was coexpressed from the plasmid encoding DHFR) at 30°C, and growth rates were monitored over a 24 hr period (details in Extended Experimental Procedures). The cells for each time point were harvested by centrifugation, resuspended in cell

resuspension buffer (Wizard), and stored at −20°C for sequencing sample preparation.

Analysis of Sequencing Data

For three independent repeats of the selection experiment, we conducted Illumina GAIIx sequencing. For two of the three experimental repeats, the sample was resequenced on multiple lanes (see Extended Experimental Procedures; Table S6), and reads from all lanes were pooled. The sequences were filtered using the associated quality scores to ensure that the barcoded regions had a probability of less than 5% of individual base miscalls. Following this filtering step, roughly 50%–90% of the reads remained, which were then sorted by experimental condition and variant using the barcodes. Variant frequencies were determined relative to WT, and normalized to the initial frequency distribution at t = 0 as follows:

$$f(t) = \log\left(\frac{N_t^{Mut}/N_t^{WT}}{N_{t-0}^{Mut}/N_{t=0}^{WT}}\right)$$

Plotting the normalized frequencies with respect to time permits reconstruction of the divergence in growth rate between WT DHFR and the other DHFR variants (Figure 3C). The relative growth rate (γ) was defined as the slope obtained by linear regression of these data (Figure 3D).

Analysis of Light Dependence and Sector Connectivity

Lit versus dark growth rates were then compared for the set of ten nonlight-dependent point mutants (lacking a LOV2 domain), and a linear fit was performed; the lit growth rates for all DHFR variants were then scaled appropriately using the linear fit parameters (repeat 1, y = 0.9853x − 0.0029; repeat 2, 1.038x + 0.0063; repeat 3, 0.9916x + 0.0041). This normalization eliminates any systematic differences in growth rate between the lit and dark experiments. Light dependence was calculated for each experiment separately as the difference in lit versus dark growth rates ($\Delta\gamma = \gamma_{lit} - \gamma_{dark}$). The significance of light dependence was determined by comparing the mean light dependence for each chimera over the three independent growth/sequencing experiments with the distribution of growth rate differences for the ten nonlight-dependent variants. To do this, we calculated a standard Z score for the light dependence of each chimera *i*: $Z_i = (\overline{\Delta\gamma_i} - \overline{\Delta\gamma_{LI}})/\sqrt{\sigma_i^2/n + \sigma_{LI}^2}$, where $\overline{\Delta\gamma_i}$ is the mean over $n = 3$ independent measurements of $\Delta\gamma$ and σ_i is the standard deviation (σ). Note that $\overline{\Delta\gamma_{LI}}$ is the mean $\overline{\Delta\gamma}$ for the ten light-independent controls (Figure 5C), and thus, σ_{LI} is a SEM. Z_i cutoffs between 2.0 and 2.8 were used for analysis, as described in the text.

Sector connectivity of each LOV2 insertion site was assessed by quantitative analysis of contacts between atoms in a high-resolution crystal structure of *E. coli* DHFR (PDB 1RX2). For each insertion site a sector contact is defined if any atom of a sector residue occurs within a specific distance cutoff from the peptide bond atoms (backbone N, Cα, C, and O) corresponding to each insertion site. Two approaches for distance cutoff from backbone atoms were examined that gave identical results: (1) the sum of the Pauling radii plus 20%, and (2) a 4 Å spherical shell.

Protein Expression and Purification

DHFR point and double mutants were expressed with a histidine tag from the pHis8-3 vector in BL21 (DE3) cells grown at 37°C in 20 ml Terrific broth to an absorbance at 600 nm of ~0.7 and induced with 0.25 mM IPTG at 18°C overnight. Cell pellets were lysed by three freeze-thaw cycles in 4 ml binding buffer (0.5 M NaCl, 10 mM imidazole, 50 mM Tris-HCl [pH 8.0]) plus lysozyme followed by centrifugation and incubation with 150 μl Ni-NTA resin (QIAGEN) for 30 min at 4°C. After washing twice with binding buffer (1 ml/wash), DHFR protein was eluted with elution buffer (1 M NaCl, 400 mM imidazole, 100 mM Tris-HCl [pH 8.0]). Eluted protein was dialyzed into dialysis buffer (300 mM NaCl, 1% glycerol/50 mM Tris-HCl [pH 8.0]) at 4°C. Purified protein was concentrated and flash frozen using liquid N_2 prior to enzymatic assays.

Steady-State Kinetic Measurements

Steady-state kinetics measurements were performed as described previously (Rajagopalan et al., 2002). Purified protein (2–500 nM) was preincubated with 100 μM NADPH in MTEN buffer at pH 7.0 (50 mM MES [2-(N-morpholino)ethanesulfonic acid], 25 mM Tris [tris(hydroxymethyl)aminomethane], 25 mM ethanolamine, and 100 mM NaCl) containing 5 mM DTT, and the reaction was initiated by adding DHF. The DHF concentration was varied according to the K_m of each enzyme but generally spanned a range from 1 to 100 μM (Figure S2). The decrease in absorbance was monitored at 340 nm ($\Delta\varepsilon_{340}$ = 13.2 mM^{-1} cm^{-1}) for at least 2 min using a Lambda 18 spectrophotometer (PerkinElmer). The steady-state parameter k_{cat} was determined under saturating concentrations of DHF (100 μM). Enzyme concentrations were determined by absorbance at 280 nm using extinction coefficients estimated by the ProtParam web server (http://ca.expasy.org/tools/protparam.html) (ε_{280}(WT, L54I, G121V, D27N,M42F, F31V, L54I/G121V, M42F/G121V) = 33585; ε_{280}(DHFR/LOV2_121, DHFR/LOV2_121-C450S) = 49055; ε_{280}(W22H) = 28085; ε_{280}(F31Y, F31Y/G121V, F31Y/L54I) = 35075).

PDZ Domain Mutagenesis and Functional Assay

Methods for the global mutational study of PDZ domains will be described in full elsewhere. Briefly, $PSD95^{pdz3}$ mutant libraries were created using a degenerate oligonucleotide-based protocol, and were assayed in a quantitative bacterial two-hybrid system in which the expression of green fluorescent protein (GFP) is proportional to the binding free energy between $PSD95^{pdz3}$ and its cognate carboxy-terminal peptide ligand derived from the CRIPT protein. Bacterial cells carrying the complete library of surface site mutations were sorted on a flow cytometer for those expressing above a threshold amount of GFP. Plasmid DNA was isolated from both sorted and unsorted bacterial populations and subject to Solexa-based high-throughput amplicon sequencing to count the frequency of observing each mutant in the two populations. For each amino acid substitution *x* at each position *i* we derive a parameter $E_i^x = \log[f_i^{x,sel}/f_i^{x,unsel}] - \log[f_i^{wt,sel}/f_i^{wt,unsel}]$, which quantitatively gives the relative fitness of that allele relative to WT. These data are shown in matrix format (Figure S7A) and used to define functionally significant sites (Figures 7 and S7B).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and six tables and can be found with this article online at doi:10.1016/j.cell.2011.10.049.

ACKNOWLEDGMENTS

We thank P.E. Wright for sharing data and discussion. We are also appreciative to S.J. Benkovic for discussions of DHFR mechanism and providing the *E. coli* auxotroph strain, M. Socolich for technical assistance, and members of the R.R. laboratory for critical review of the work. This study was supported by the Robert A. Welch foundation (R.R., I-1366), the NIH (1RO1 EY018720), and the Green Center for Systems Biology at UT Southwestern Medical Center.

Received: April 28, 2011
Revised: August 10, 2011
Accepted: October 19, 2011
Published: December 22, 2011

REFERENCES

Agarwal, P.K., Billeter, S.R., Rajagopalan, P.T., Benkovic, S.J., and Hammes-Schiffer, S. (2002). Network of coupled promoting motions in enzyme catalysis. Proc. Natl. Acad. Sci. USA *99*, 2794–2799.

Almasan, A., Yin, Y., Kelly, R.E., Lee, E.Y., Bradley, A., Li, W., Bertino, J.R., and Wahl, G.M. (1995). Deficiency of retinoblastoma protein leads to inappropriate S-phase entry, activation of E2F-responsive genes, and apoptosis. Proc. Natl. Acad. Sci. USA *92*, 5436–5440.

Benkovic, S.J., and Hammes-Schiffer, S. (2003). A perspective on enzyme catalysis. Science *301*, 1196–1202.

Bhabha, G., Lee, J., Ekiert, D.C., Gam, J., Wilson, I.A., Dyson, H.J., Benkovic, S.J., and Wright, P.E. (2011). A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis. Science *332*, 234–238.

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

Boehr, D.D., McElheny, D., Dyson, H.J., and Wright, P.E. (2010). Millisecond timescale fluctuations in dihydrofolate reductase are exquisitely sensitive to the bound ligands. Proc. Natl. Acad. Sci. USA *107*, 1373–1378.

Breslow, D.K., Cameron, D.M., Collins, S.R., Schuldiner, M., Stewart-Ornstein, J., Newman, H.W., Braun, S., Madhani, H.D., Krogan, N.J., and Weissman, J.S. (2008). A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. Nat. Methods *5*, 711–718.

Chen, J., Dima, R.I., and Thirumalai, D. (2007). Allosteric communication in dihydrofolate reductase: signaling network and pathways for closed to occluded transition and back. J. Mol. Biol. *374*, 250–266.

Clarkson, M.W., Gilmore, S.A., Edgell, M.H., and Lee, A.L. (2006). Dynamic coupling and allosteric behavior in a nonallosteric protein. Biochemistry *45*, 7693–7699.

Cui, Q., and Karplus, M. (2008). Allostery and cooperativity revisited. Protein Sci. *17*, 1295–1307.

Eisenmesser, E.Z., Millet, O., Labeikovsky, W., Korzhnev, D.M., Wolf-Watz, M., Bosco, D.A., Skalicky, J.J., Kay, L.E., and Kern, D. (2005). Intrinsic dynamics of an enzyme underlies catalysis. Nature *438*, 117–121.

Ferguson, A.D., Amezcua, C.A., Halabi, N.M., Chelliah, Y., Rosen, M.K., Ranganathan, R., and Deisenhofer, J. (2007). Signal transduction pathway of TonB-dependent transporters. Proc. Natl. Acad. Sci. USA *104*, 513–518.

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

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

Gether, U. (2000). Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. Endocr. Rev. *21*, 90–113.

Halabi, N., Rivoire, O., Leibler, S., and Ranganathan, R. (2009). Protein sectors: evolutionary units of three-dimensional structure. Cell *138*, 774–786.

Halavaty, A.S., and Moffat, K. (2007). N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from *Avena sativa*. Biochemistry *46*, 14001–14009.

Harper, S.M., Neil, L.C., and Gardner, K.H. (2003). Structural basis of a phototropin light switch. Science *301*, 1541–1544.

Hatley, M.E., Lockless, S.W., Gibson, S.K., Gilman, A.G., and Ranganathan, R. (2003). Allosteric determinants in guanine nucleotide-binding proteins. Proc. Natl. Acad. Sci. USA *100*, 14445–14450.

Kuriyan, J., and Eisenberg, D. (2007). The origin of protein interactions and allostery in colocalization. Nature *450*, 983–990.

Lee, J., Natarajan, M., Nashine, V.C., Socolich, M., Vo, T., Russ, W.P., Benkovic, S.J., and Ranganathan, R. (2008). Surface sites for engineering allosteric control in proteins. Science *322*, 438–442.

Lee, S.Y., Banerjee, A., and MacKinnon, R. (2009). Two separate interfaces between the voltage sensor and pore are required for the function of voltage-dependent K(+) channels. PLoS Biol. *7*, e47.

Lockless, S.W., and Ranganathan, R. (1999). Evolutionarily conserved pathways of energetic connectivity in protein families. Science *286*, 295–299.

Luque, I., Leavitt, S.A., and Freire, E. (2002). The linkage between protein folding and functional cooperativity: two sides of the same coin? Annu. Rev. Biophys. Biomol. Struct. *31*, 235–256.

Marcotte, E.M., Pellegrini, M., Ng, H.L., Rice, D.W., Yeates, T.O., and Eisenberg, D. (1999). Detecting protein function and protein-protein interactions from genome sequences. Science *285*, 751–753.

McElheny, D., Schnell, J.R., Lansing, J.C., Dyson, H.J., and Wright, P.E. (2005). Defining the role of active-site loop fluctuations in dihydrofolate reductase catalysis. Proc. Natl. Acad. Sci. USA *102*, 5032–5037.

Menon, S.T., Han, M., and Sakmar, T.P. (2001). Rhodopsin: structural basis of molecular physiology. Physiol. Rev. *81*, 1659–1688.

Metzker, M.L. (2010). Sequencing technologies—the next generation. Nat. Rev. Genet. *11*, 31–46.

Monod, J., Wyman, J., and Changeux, J.P. (1965). On the nature of allosteric transitions: a plausible model. J. Mol. Biol. *12*, 88–118.

Niethammer, M., Valtschanoff, J.G., Kapoor, T.M., Allison, D.W., Weinberg, R.J., Craig, A.M., and Sheng, M. (1998). CRIPT, a novel postsynaptic protein that binds to the third PDZ domain of PSD-95/SAP90. Neuron *20*, 693–707.

Nourry, C., Grant, S.G., and Borg, J.P. (2003). PDZ domain proteins: plug and play!. Sci. STKE *2003*, RE7.

Peterson, F.C., Penkert, R.R., Volkman, B.F., and Prehoda, K.E. (2004). Cdc42 regulates the Par-6 PDZ domain through an allosteric CRIB-PDZ transition. Mol. Cell *13*, 665–676.

Rajagopalan, P.T., and Benkovic, S.J. (2002). Preorganization and protein dynamics in enzyme catalysis. Chem. Rec. *2*, 24–36.

Rajagopalan, P.T., Lutz, S., and Benkovic, S.J. (2002). Coupling interactions of distal residues enhance dihydrofolate reductase catalysis: mutational effects on hydride transfer rates. Biochemistry *41*, 12618–12628.

Russ, W.P., Lowery, D.M., Mishra, P., Yaffe, M.B., and Ranganathan, R. (2005). Natural-like function in artificial WW domains. Nature *437*, 579–583.

Sadovsky, E., and Yifrach, O. (2007). Principles underlying energetic coupling along an allosteric communication trajectory of a voltage-activated K+ channel. Proc. Natl. Acad. Sci. USA *104*, 19813–19818.

Salomon, M., Eisenreich, W., Dürr, H., Schleicher, E., Knieb, E., Massey, V., Rüdiger, W., Müller, F., Bacher, A., and Richter, G. (2001). An optomechanical transducer in the blue light receptor phototropin from *Avena sativa*. Proc. Natl. Acad. Sci. USA *98*, 12357–12361.

Saraf, M.C., Horswill, A.R., Benkovic, S.J., and Maranas, C.D. (2004). FamClash: a method for ranking the activity of engineered enzymes. Proc. Natl. Acad. Sci. USA *101*, 4142–4147.

Schnell, J.R., Dyson, H.J., and Wright, P.E. (2004). Structure, dynamics, and catalytic function of dihydrofolate reductase. Annu. Rev. Biophys. Biomol. Struct. *33*, 119–140.

Shulman, A.I., Larson, C., Mangelsdorf, D.J., and Ranganathan, R. (2004). Structural determinants of allosteric ligand activation in RXR heterodimers. Cell *116*, 417–429.

Smock, R.G., and Gierasch, L.M. (2009). Sending signals dynamically. Science *324*, 198–203.

Smock, R.G., Rivoire, O., Russ, W.P., Swain, J.F., Leibler, S., Ranganathan, R., and Gierasch, L.M. (2010). An interdomain sector mediating allostery in Hsp70 molecular chaperones. Mol. Syst. Biol. *6*, 414.

Socolich, M., Lockless, S.W., Russ, W.P., Lee, H., Gardner, K.H., and Ranganathan, R. (2005). Evolutionary information for specifying a protein fold. Nature *437*, 512–518.

Süel, G.M., Lockless, S.W., Wall, M.A., and Ranganathan, R. (2003). Evolutionarily conserved networks of residues mediate allosteric communication in proteins. Nat. Struct. Biol. *10*, 59–69.

Thompson, D.A., Desai, M.M., and Murray, A.W. (2006). Ploidy controls the success of mutators and nature of mutations during budding yeast evolution. Curr. Biol. *16*, 1581–1590.

Yifrach, O., and MacKinnon, R. (2002). Energetics of pore opening in a voltage-gated K(+) channel. Cell *111*, 231–239.