

REVIEW

Engineering allosteric protein switches by domain insertion

Marc Ostermeier

Department of Chemical and Biomolecular Engineering, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA

E-mail: oster@jhu.edu

Domain insertion is proving to be an effective way to construct hybrid proteins exhibiting switch-like behavior. In this strategy, two existing domains, the first exhibiting a signal recognition function and the second containing the function to be modulated, are fused such that the recognition of the signal by the first domain is transmitted to the second domain, thereby modulating its activity. Recent directed evolution experiments indicate that the structural space comprised of the recombination of unrelated protein domains may be rich in switching behavior, particularly when the circular permutation of domains is also employed. This bodes well for potential basic science, sensing and therapeutic applications of molecular switches.

Keywords: allosteric protein switches/domain insertion/molecular switches

Introduction

Biological complexity arises not only from the large number of cellular components but also from the myriad of functional relationships between these components. Key elements in this web of relationships are biomolecules that function as switches. A switch recognizes an input signal (e.g. ligand concentration, pH, covalent modification) and, as a result, its output signal (e.g. enzyme activity, ligand affinity, oligomeric state) is modulated. Typically switches have spatially distinct regulatory and active sites. Binding or modifications at one site affects the output function at a distant site, often through a conformational change. Thus, switches establish functional relationships between different biological molecules and between biological molecules and their environment. This coupling of function enables regulatory and control mechanisms central to biological complexity and to life itself.

The ability to create novel switches or to modify existing switches to create new functional relationships would enable a wide variety of applications. Switches can be envisioned that sense the cellular state and carry out functions accordingly in response to that state (e.g. conditionally deliver drugs, activate transcription, inhibit protein function, modulate signal transduction pathways or adopt a toxic conformation). In addition, the ability to create protein switches has tremendous practical potential for developing novel molecular sensors and as a tool for elucidating molecular and cellular function. The importance of controlling cellular protein function in general has inspired a wide variety of approaches—many involving reprogrammed

natural switches—that are outside the scope of this review. Recent reviews cover many of these approaches (Villaverde, 2003; Dueber *et al.*, 2004; Buskirk and Liu, 2005). This review concerns strategies for creating switches involving the insertion of one domain into another. At the core of these strategies is the concept that protein switches can be created by combining the genes of existing protein domains with the prerequisite input and output functions for the switch (Figure 1) such that changes in one domain (e.g. conformational change upon ligand binding) will modulate the function of the second domain.

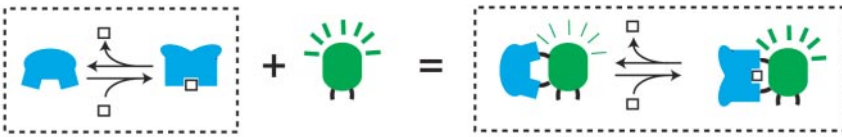
Domain insertion

The fusion of two or more genes into a single gene has been widely used as a tool in protein engineering, localization and purification. Although end-to-end fusion has been used almost exclusively, insertional fusion—in which one gene is inserted into the middle of the other gene—is an interesting alternative (Doi and Yanagawa, 1999a). For fusion by insertion, success seemingly requires that the inserted protein has its N- and C-terminal proximal and that the target domain be able to be converted from a continuous to a discontinuous domain. However, these conditions are not as restrictive as one might expect. First, a recent survey of the Protein Data Bank indicates that 50% of all single domain proteins have their N- and C-termini proximal (≤ 5 Å apart) (Krishna and Englander, 2005); hence many proteins are at least geometrically amenable to being inserted. Second, discontinuous domains (domains in which the linear sequence of the domain is interrupted by another inserted domain) are fairly prevalent in nature. A systematic survey of structural domains indicated that 28% of structural domains are discontinuous (Jones *et al.*, 1998). Accordingly, since the first example of successful insertion of one protein into another [alkaline phosphatase into the *Escherichia coli* outer membrane protein MalF, constructed as a tool for studying membrane topology (Ehrmann *et al.*, 1990)], a wide variety of protein pairs have been shown to be amenable to domain insertion fusion. The diversity in inserted sequences (Starzyk *et al.*, 1989; Doi *et al.*, 1997) and in the sites for insertion (including sites within secondary structure elements) (Barany, 1985; Biondi *et al.*, 1998; Guntas and Ostermeier, 2004), indicates that proteins can be fairly tolerant to domain insertion. Since domain insertion creates a more intimate connection between the two proteins, several groups have successfully employed this strategy to couple the functions of two proteins and create switches (Table I).

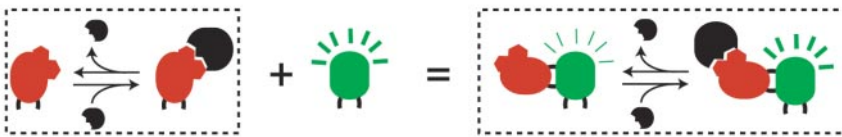
Fluorescent sensors

The optical signal transduction of the green fluorescent protein (GFP) has made it an attractive target for engineering

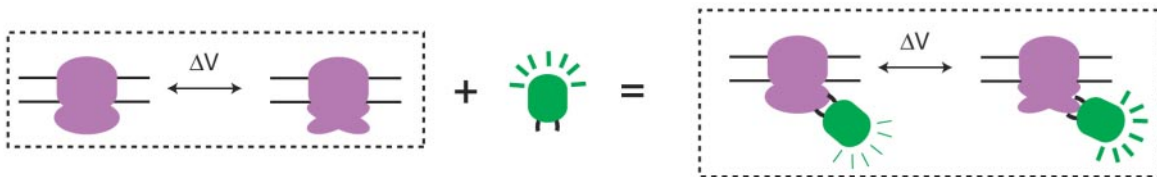
A Fluorescent sensors for ligands



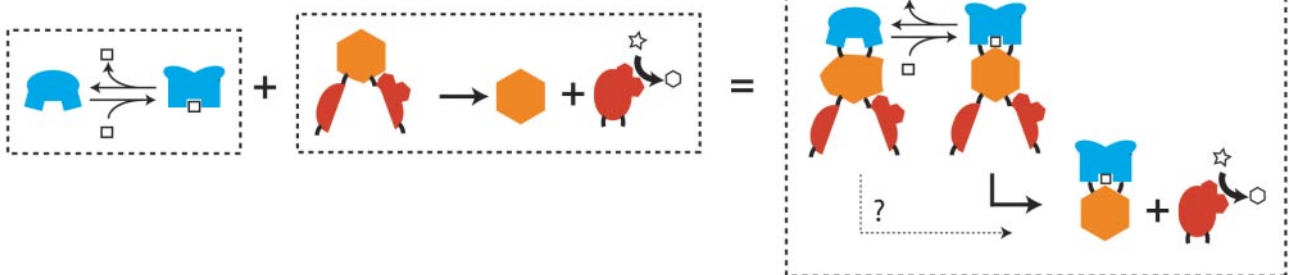
B Fluorescent sensors for proteins



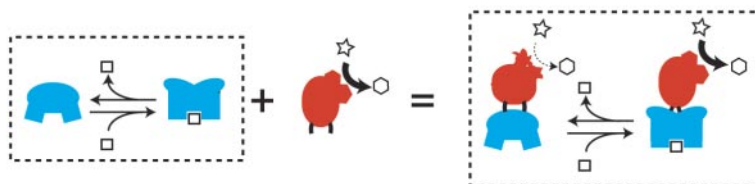
C Fluorescent sensors for transmembrane voltage changes



D Conditional protein splicing



E Allosteric enzymes



F Mutually exclusive folding



Fig. 1. Schematic depiction of several protein switches created by domain insertion. The schematics are for illustrative purposes only and are not meant to imply the relative orientation of the domains or the mechanism by which switching is occurring. **(A)** Fluorescent sensor for a small molecule ligand (open square) created by fusing a fluorescent protein (green) and a ligand-binding domain (blue) (Baird *et al.*, 1999). In the switch, ligand binding increases fluorescence. **(B)** Similar scheme in which a fluorescent sensor for an inhibitory protein (black) was created by fusing a fluorescent protein (green) with the enzyme (red) that the inhibitory protein binds (Doi and Yanagawa, 1999b). **(C)** Transmembrane voltage sensor created by inserting a fluorescent protein (green) into a voltage-activated ion channel (purple) that undergoes a conformational change upon a change in membrane potential (Siegel and Isacoff, 1997; Ataka and Pieribone, 2002). **(D)** Small molecule-dependent inteins that are tripartite fusions consisting of a ligand binding domain (blue), an intein (orange) and a protein that is to be spliced (red). Binding of the ligand (open square) activates the splicing and produces the correctly spliced, active protein (Buskirk *et al.*, 2004; Skretas and Wood, 2005). **(E)** An enzymatic switch created by fusing an enzyme (red) and a ligand binding domain (blue) that binds a small molecule ligand (open square) (Guntas and Ostermeier, 2004; Guntas *et al.*, 2004, 2005). The switch's behavior is analogous to that of a heterotropic allosteric enzyme. **(F)** A molecular switch created by inserting a protein with distal N- and C-termini (black) into another protein (gray) such that both proteins cannot exist in a properly folded conformation at the same time (Radley *et al.*, 2003). If one protein is more stable under one set of conditions and the other is more stable under another set of conditions, then switching can occur by changing between conditions (e.g. by changing temperature).

Table 1. Selected examples of molecular switches created by domain insertion^a

Host	Insert	Mut. ^b	Effector	Property modulated <i>in vitro</i>	Max. fold <i>in vitro</i> effect ^c	Phenotype modulated <i>in vivo</i>	Max. fold <i>in vivo</i> effect ^c	Reference
GFP	BLA	Yes	BLIP	Fluorescence	1.6			Doi and Yanagawa (1999b)
GFP	Calmodulin	No	Ca ²⁺	Fluorescence	7	Fluorescence	7	Baird <i>et al.</i> (1999)
Ion channel	GFP	No	ΔV			Fluorescence	1.05	Siegel and Isacoff (1997)
DHFR	FKBP-12	Yes	FK506			Density of yeast cells	2	Tucker and Fields (2001)
DHFR	Estrogen receptor α	No	17 β -Estradiol			Density of yeast cells	2.5	Tucker and Fields (2001)
RecA intein	Estrogen receptor α	Yes	4-Hydroxytamoxifen			Phenotype associated with spliced product	$\geq 10^d$	Buskirk <i>et al.</i> (2004)
Δ I-SM mini-intein	Thyroid receptor β	No	Thyroid hormone T3			Phenotype associated with spliced product	4 ^e	Skretas and Wood (2005)
Δ I-SM mini-intein	Estrogen receptor α	Yes	Diethylstilbestrol			Phenotype associated with spliced product	2 ^e	Skretas and Wood (2005)
MBP	BLA	No	Maltose	Rate of β -lactam hydrolysis	600	<i>E.coli</i> MIC ampicillin	16	Guntas <i>et al.</i> (2005)
MBP	BLA	Yes ^f	Sucrose	Rate of β -lactam hydrolysis	82	<i>E.coli</i> MIC ampicillin	2	Guntas <i>et al.</i> (2005)

^aBLA, TEM1 β -lactamase; BLIP, β -lactamase inhibitory protein; DHFR, dihydrofolate reductase; FKBP-12, FK506 binding protein; GFP, green fluorescent protein; MBP, maltose binding protein; ΔV , change in membrane potential.

^bDid the switch require mutations in either host or insert domain for switching activity?

^cValue of property/phenotype in the presence of the effector divided by the value in the absence of the effector.

^dProduction of LacZ enzymatic activity.

^eDensity of *E.coli* culture.

^fMutations required in order to bind sucrose.

biosensors by domain insertion. The most impressive fluorescent sensor created by domain insertion is that of *Xenopus* calmodulin (a Ca²⁺-binding protein) inserted into a variant of GFP (Figure 1A) (Baird *et al.*, 1999). The calmodulin domain undergoes a large conformational change upon binding of Ca²⁺ that, in the fusion, results in a 7-fold increase in the fluorescence intensity in the presence of saturating amounts of calcium. Evidence suggests that the conformational change results in a change in the protonation state of the fluorophore, thereby increasing its fluorescence. The sensors have been successfully used to detect intracellular calcium dynamics in mammalian cells (Baird *et al.*, 1999; Griesbeck *et al.*, 2001) and the brain of *Drosophila* (Yu *et al.*, 2003).

GFP sensors created by domain insertion do not necessarily require a binding event with as dramatic a conformational change as calmodulin. Although initial attempts at creating a molecular sensor for the β -lactamase inhibitory protein (BLIP) by inserting β -lactamase into GFP were unsuccessful, random mutagenesis on the gene fusion was able to create a protein whose fluorescence increased by 60% upon BLIP binding (Figure 1B) (Doi and Yanagawa, 1999b). In addition, GFP has been inserted into voltage-gated channels for potassium (Siegel and Isacoff, 1997) and sodium (Ataka and Pieribone, 2002) to generate sensors in which voltage-driven rearrangements in the channel alter the brightness of GFP by 5.1 and 0.5%, respectively (Figure 1C).

Small molecule-dependent protein splicing

Inteins are the protein equivalent of introns and exist as in-frame insertions within the host protein. Typically, the host protein is inactivated by this insertion. Inteins possess the ability to ligate their flanking sequences and excise themselves from their protein host. Independently, two groups have engineered small molecule-dependent inteins by insertion of

hormone binding domains into an intein, which in turn has been inserted into the protein to be spliced (Figure 1D) (Buskirk *et al.*, 2004; Skretas and Wood, 2005). In these switches, the efficiency of protein splicing is linked to the presence of the hormone. These intein switches are thus methods of controlling the production of a protein at the protein level. The generality of the approach—the fact that a single chemical signal can be easily adapted to controlling the production of many different proteins—is a clear strength of these switches and has obvious parallels with and advantages over transcriptional and translational controls of protein production. However, from the perspective of controlling biological function, the benefit of generality comes at the expense of slower dynamics and not being able to control the spliced protein once produced.

Allosteric enzymes

Since domain insertions result in a continuous domain being split into a discontinuous domain, such fusions can be used to address important issues in protein folding. Collinet *et al.* (2000) constructed several domain insertion fusions between different enzymes for this purpose. Most interestingly, for some of the constructs, the catalytic rate of one enzyme domain in the fusion was affected by the presence of the substrate for the other domain. Although the effect was small (at most a 1.6-fold change) and not extensively characterized, the work suggested the sensitivity of enzyme domains in domain insertion fusions.

We have constructed a family of enzymatic switches by recombining the *E.coli* maltose binding protein (MBP) and TEM1 β -lactamase (BLA) such that maltodextrins are either a positive or negative effector of BLA enzyme activity (Figure 1E). We accomplished this by combinatorially exploring a structural space that can be conceptualized as ‘rolling’ the two proteins across each others’ surface and fusing them

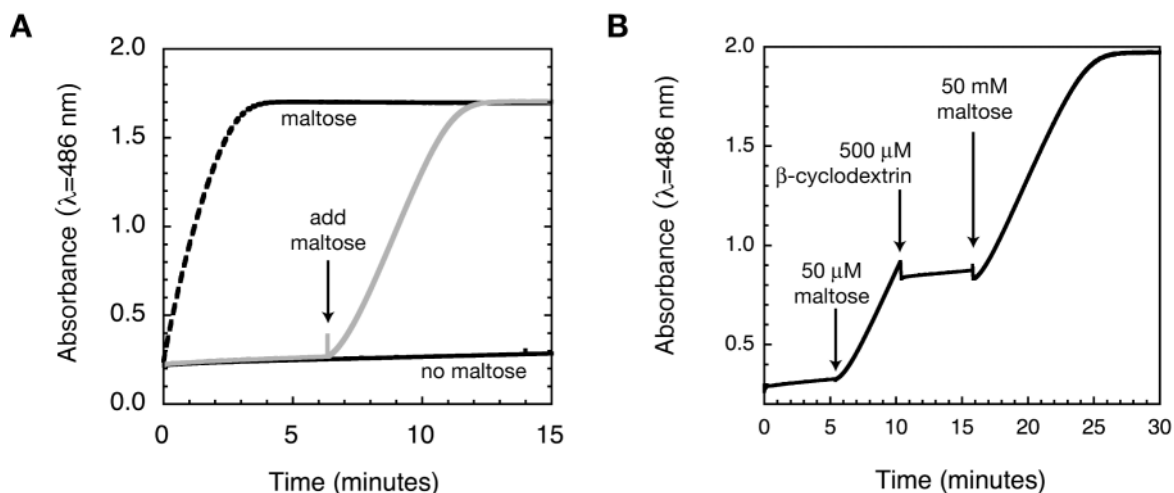


Fig. 2. Maltose activation of BLA activity of MBP-BLA switch MBP317-347. During the enzymatic hydrolysis of the BLA substrate nitrocefin, formation of product is monitored by measuring the absorbance at 486 nm. **(A)** Hydrolysis of nitrocefin by MBP317-347 in the presence and absence of maltose at 25°C (Guntas *et al.*, 2005). The reaction was started by the addition of nitrocefin at time zero to samples lacking (solid lines) or containing (dashed line) 5 mM maltose. For the reaction traced by the solid gray line, 5 mM maltose was added to the reaction at about 6 min. **(B)** Reversible switching using competing ligand. At time zero the reaction was started in the absence of any effector. At the time indicated by the first arrow, maltose was added. This maltose concentration was above the K_d for maltose but was subsaturating. At the time indicated by the second arrow, β -cyclodextrin was added in 10-fold excess of maltose. β -Cyclodextrin binds to MBP317-347 at the same site as maltose but does not activate it. Because MBP317-347 has similar affinity for maltose and β -cyclodextrin, the β -cyclodextrin (in excess) preferentially replaced the maltose bound to MBP317-347 and the switch was turned back off. At the time indicated by the third arrow, an excess of maltose was added and the switch was turned back on and the reaction was allowed to go to completion.

through peptide bonds at the points where their surfaces meet. We first demonstrated that one could randomly insert the BLA gene into the MBP gene and, through selection/screening strategies, create switches (Guntas and Ostermeier, 2004). Although the results were significant in the sense that the method created a new property (allostery) that did not exist before, the switching effect was modest (<2-fold). We next explored a more radical recombination of the two genes: simultaneous random circular permutation (Graf and Schachman, 1996) of the BLA gene followed by random insertion of the circularly permuted library into the MBP gene. This approach was more successful and a switch was identified in which maltose-binding increased BLA catalytic efficiency 25-fold (Guntas *et al.*, 2004). Subsequently, we used an iterative approach involving alternating random domain insertion and random circular permutation to create several switches that approach what can effectively be described as being ‘on/off’ in nature (maltose increased BLA activity several hundred-fold, see Figure 2A) (Guntas *et al.*, 2005). In addition to the unprecedented level of switching, these fusion proteins exhibited a number of desirable properties including specificity (they only responded to known ligands of MBP), dose–response-type behavior and reversibility (upon removal of maltose the BLA rate decreased back to the low level; Figure 2B). The switches conferred upon *E. coli* a novel phenotype: maltose-dependent resistance to β -lactam antibiotics. This illustrated two key attributes of switches: (1) their ability to ‘rewire’ the cell (maltose concentrations and antibiotic resistance are coupled) and (2) their potential for sensing applications (the *E. coli* cells are growth/no-growth ‘biosensors’ for maltose).

We have shown that existing MBP-BLA switches can be made to respond to lower levels of maltose by incorporating mutations known to increase MBP affinity for maltose (Guntas *et al.*, 2004). In addition, we have engineered sucrose-dependent switches by exploiting the ligand-dependent antibiotic resistance conferred by the MBP-BLA switches (Guntas

et al., 2005). The binding pocket of the MBP domain in the MBP-BLA switch was randomized and cells bearing these mutated switches were plated in the presence of sucrose at concentrations of β -lactam antibiotic that normally only allow growth if maltose is present. Under these selective conditions, a library member that is able to bind sucrose and trigger the switch to adopt its more active conformation for BLA activity will confer a high level of β -lactam antibiotic resistance and allow growth. Using this selection scheme we created a switch that responds to sucrose as the effector and bound sucrose with an affinity equivalent to that of MBP for maltose.

Cellular sensors of ligand binding

Tucker and Fields (2001) based their approach for creating switches on the principle of ligand-induced stabilization. Two types of yeast cells capable of having a ligand-dependent growth phenotype were created via domain insertion fusions. FKBP-12 was inserted into a temperature-sensitive mutant of dihydrofolate reductase (DHFR) and ER α was inserted into wild-type DHFR (in this case, the act of insertion itself created temperature-sensitive DHFR activity). Yeast grown under conditions in which DHFR activity was limiting the growth rate exhibited a ligand-dependent growth rate and final cell density. The DHFR activity of the purified fusions did not exhibit significant changes in stability or activity in the presence of ligand *in vitro*, leading to speculation that very subtle changes can have a large effect *in vivo* or that the ligand bound form may be more proteolytically stable.

Mutually exclusive folding

Mutually exclusive folding is conceptually different from the preceding examples in that the inserted protein has its N- and C-termini far away from each other by design. This geometric constraint prevents both domains from being folded at the same

time (Figure 1F). Loh and co-workers reported such a system in which human ubiquitin was inserted into barnase (Radley *et al.*, 2003). At low temperatures, barnase is more stable and its folded conformation prevents ubiquitin from adopting its properly folded conformation. At elevated temperatures, ubiquitin is more stable and therefore folded, whereas the barnase domain is unfolded. In addition to using temperature as a method to switch between two activities, ligands that increase stability can be used. Barstar is a protein inhibitor of barnase that stabilizes the folded conformation of barnase. Thus, at a temperature in which ubiquitin is marginally more stable, barstar can be used as an effector to switch the conformational equilibrium toward having folded barnase.

Mechanism of switching

For switches created by domain insertion, the functional relationship between the input and output domains is probably complex and similar to that of conventional allosteric proteins in which allostery is achieved through a chain of propagating interactions or a shift in the conformational ensemble (Luque and Freire, 2000; Pan *et al.*, 2000; Yu and Koshland, 2001; Kern and Zuiderweg, 2003; Suel *et al.*, 2003; Gunasekaran *et al.*, 2004). However, such a view is speculative since the mechanism for switching has not been determined for any switch created by domain insertion. For our MBP-BLA switches, the switching depends on a conformational change in the MBP domain of the switch, as might be expected (Guntas and Ostermeier, 2004; Guntas *et al.*, 2004). Mutations known to shift the *apo* form of MBP to a more closed state, thereby increasing maltose affinity, also increase maltose affinity in the switch. More importantly, mutations that close the *apo* form of MBP the furthest abrogate switching by raising the BLA activity of the switch in the absence of maltose (Guntas *et al.*, 2004). Although a conformational change in the MBP domain is implicated by these studies, how the conformational change in MBP affects BLA is not clear. However, the complex nature of the switching observed (i.e. maltose has different effects on different BLA substrates) seemingly rules out mechanisms involving steric occlusion or altering the equilibrium between a folded and unfolded state. We speculate that the BLA active site may be 'forcibly held' in a reorganized, generally less active conformation (or ensemble of conformations) in the absence of maltose, a constraint that is relaxed upon maltose binding.

Outlook

In the most optimistic view, creating switches by domain insertion is limited only by the availability of domains with the prerequisite functions. Given the accelerated rate that genomic and proteomic efforts are providing us with a 'biological parts list', this is perhaps not too serious a limitation. An open question, however, is how often domain insertion or other methods of recombination of these parts can result in switching behavior. The author is optimistic that establishing functional relationships between two proteins by domain insertion will prove to be relatively easy to accomplish, at least for low-level switching. Several studies were successful in creating switches by trial-and-error and we have identified a wide variety of MBP-BLA switches through our combinatorial approach despite examining only a small fraction of the possible

recombinations between the two genes. Even for fusions that do not behave as switches, several examples show how switching activity can be easily evolved by random mutagenesis (Doi and Yanagawa, 1999b; Buskirk *et al.*, 2004; Skretas and Wood, 2005). Despite this, computational strategies for creating switches face considerable challenges considering the difficulty in understanding the structural mechanisms of natural allosteric proteins and in predicting the structure of domain insertion fusions. Thus, in the near future, directed evolution approaches hold the most promise for creating switches with the properties desired for applications. Such an approach is fitting considering Monod *et al.*'s view that allosteric proteins are 'those most elaborate products of molecular evolution' (Monod *et al.*, 1963). Detailed mechanistic and structural studies on switches arrived at by directed evolution will be invaluable for understanding how switching properties can be engineered and may also shed light on natural allosteric mechanisms.

Acknowledgements

I thank Jeffrey J. Gray and Gurkan Guntas for helpful discussions on this review. I also thank Gurkan Guntas for performing the experiment depicted in Figure 2B. This work was supported by a grant from NIH.

References

- Ataka, K. and Pieribone, V.A. (2002) *Biophys. J.*, **82**, 509–516.
- Baird, G.S., Zacharias, D.A. and Tsien, R.Y. (1999) *Proc. Natl Acad. Sci. USA*, **96**, 11241–11246.
- Barany, F. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 4202–4206.
- Biondi, R.M., Baehler, P.J., Reymond, C.D. and Véron, M. (1998) *Nucleic Acids Res.*, **26**, 4946–4952.
- Buskirk, A.R. and Liu, D.R. (2005) *Chem. Biol.*, **12**, 151–161.
- Buskirk, A.R., Ong, Y.C., Gartner, Z.J. and Liu, D.R. (2004) *Proc. Natl Acad. Sci. USA*, **101**, 10505–10510.
- Collinet, B., Herve, M., Pecorari, F., Minard, P., Eder, O. and Desmadril, M. (2000) *J. Biol. Chem.*, **275**, 17428–17433.
- Doi, N. and Yanagawa, H. (1999a) *FEBS Lett.*, **457**, 1–4.
- Doi, N. and Yanagawa, H. (1999b) *FEBS Lett.*, **453**, 305–307.
- Doi, N., Itaya, M., Yomo, T., Tokura, S. and Yanagawa, H. (1997) *FEBS Lett.*, **402**, 177–180.
- Dueber, J.E., Yeh, B.J., Bhattacharyya, R.P. and Lim, W.A. (2004) *Curr. Opin. Struct. Biol.*, **14**, 690–699.
- Ehrmann, M., Boyd, D. and Beckwith, J. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 7574–7578.
- Graf, R. and Schachman, H.K. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 11591–11596.
- Griesbeck, O., Baird, G.S., Campbell, R.E., Zacharias, D.A. and Tsien, R.Y. (2001) *J. Biol. Chem.*, **276**, 29188–29194.
- Gunasekaran, K., Ma, B. and Nussinov, R. (2004) *Proteins*, **57**, 433–443.
- Guntas, G. and Ostermeier, M. (2004) *J. Mol. Biol.*, **336**, 263–273.
- Guntas, G., Mitchell, S.F. and Ostermeier, M. (2004) *Chem. Biol.*, **11**, 1483–1487.
- Guntas, G., Mansell, T., Kim, J.R. and Ostermeier, M. (2005) *Proc. Natl Acad. Sci. USA*, in press.
- Jones, S., Stewart, M., Michie, A., Swindells, M.B., Orengo, C. and Thornton, J.M. (1998) *Protein Sci.*, **7**, 233–242.
- Kern, D. and Zuiderweg, E.R. (2003) *Curr. Opin. Struct. Biol.*, **13**, 748–757.
- Krishna, M.M. and Englander, S.W. (2005) *Proc. Natl Acad. Sci. USA*, **102**, 1053–1058.
- Luque, I. and Freire, E. (2000) *Proteins*, Suppl., 63–71.
- Monod, J., Changeux, J.P. and Jacob, F. (1963) *J. Mol. Biol.*, **6**, 306–329.
- Pan, H., Lee, J.C. and Hilser, V.J. (2000) *Proc. Natl Acad. Sci. USA*, **97**, 12020–12025.
- Radley, T.L., Markowska, A.I., Bettinger, B.T., Ha, J.H. and Loh, S.N. (2003) *J. Mol. Biol.*, **332**, 529–536.
- Siegel, M.S. and Isacoff, E.Y. (1997) *Neuron*, **19**, 735–741.
- Skretas, G. and Wood, D.W. (2005) *Protein Sci.*, **14**, 523–352.
- Starzyk, R.M., Burbaum, J.J. and Schimmel, P. (1989) *Biochemistry*, **28**, 8479–8484.
- Suel, G.M., Lockless, S.W., Wall, M.A. and Ranganathan, R. (2003) *Nat. Struct. Biol.*, **10**, 59–69.

M.Ostermeier

Tucker,C.L. and Fields,S. (2001) *Nat. Biotechnol.*, **19**, 1042–1046.

Villaverde,A. (2003) *FEBS Lett.*, **554**, 169–172.

Yu,D., Baird,G.S., Tsien,R.Y. and Davis,R.L. (2003) *J. Neurosci.*, **23**, 64–72.

Yu,E.W. and Koshland,D.E.,Jr (2001) *Proc. Natl Acad. Sci. USA*, **98**, 9517–9520.

Received May 20, 2005; revised July 1, 2005; accepted July 4, 2005

Edited by Dan Tawfik