Directed evolution of protein switches and their application to the creation of ligand-binding proteins

Gurkan Guntas*, Thomas J. Mansell, Jin Ryoun Kim, and Marc Ostermeier*†

Department of Chemical and Biomolecular Engineering, The Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218

Edited by Stanley Fields, University of Washington, Seattle, WA, and approved June 24, 2005 (received for review March 31, 2005)

We describe an iterative approach for creating protein switches involving the *in vitro* recombination of two nonhomologous genes. We demonstrate this approach by recombining the genes coding for TEM1 β-lactamase (BLA) and the *Escherichia coli* maltose binding protein (MBP) to create a family of MBP–BLA hybrids in which maltose is a positive or negative effector of β-lactam hydrolysis. Some of these MBP–BLA switches were effectively "on–off" in nature, with maltose altering catalytic activity by as much as 600-fold. The ability of these switches to confer an effector-dependent growth/no growth phenotype to *E. coli* cells was exploited to rapidly identify, from a library of 4×10^6 variants, MBP–BLA switch variants that respond to sucrose as the effector. The transplantation of these mutations into wild-type MBP converted MBP into a "sucrose-binding protein," illustrating the switches potential as a tool to rapidly identify ligand-binding proteins.

allostery | β-lactamase | maltose binding protein

Regulation of protein activity is fundamental to cellular function. One of the mechanisms a cell uses to modulate the level of protein activity is regulation of the amount of a protein present in a cell. Accordingly, many strategies for engineering control of cellular protein activity have focused on modulation of protein production and degradation, often through small-molecule-dependent switches that regulate transcription, translation, localization, degradation, or protein splicing (1) or through the engineering of artificial gene-regulatory networks (2). The key strength of many of these approaches is that they are easily generalized. For example, a switch that regulates transcription of one gene can easily be adapted to regulate transcription of an arbitrary gene. However, the significant limitation of these approaches is the slow dynamics stemming from the indirect nature of the regulation (i.e., protein activity is regulated by controlling the amount of protein and not by regulating the protein's specific activity directly). In addition, modulation is only feasible in the context of the cell and cannot be easily transferred to an *in vitro* setting.

A more satisfying approach is to modulate the protein activity directly, but this presents a considerable design problem. Inhibitors, if they can be found, can only be used to down-regulate activity, and such a strategy suffers from the fact that one is not free to choose the signal that modulates the activity: The signal must be an inhibitor of the protein. One clever way around this limitation is to engineer the inhibitor such that a third molecule can regulate the inhibitor's affinity for the regulated protein, as was demonstrated for RNA aptamer inhibitors that could be regulated by an organic small molecule (3). Natural allosteric proteins solve this problem by having spatially distinct regulatory and active sites. Ligand binding or covalent modifications at one site affects the output function at a distant site through a conformational change. This mechanism has inspired a modular design strategy for creating switches in which existing proteins with the prerequisite input and output functions are combined such that their functions are coupled (4–13). Such switches have potential applications as molecular sensors, as tools for metabolic and cellular engineering, and as therapeutic proteins. Before these applications are realized, the considerable design problem of exactly how to combine two unrelated proteins in a way that significantly couples their activities must be overcome.

We have recently demonstrated that a single round of nonhomologous recombination followed by genetic selection/screening can create modest switches from genes coding for the desired input and output modules (9, 12). We recombined the genes coding for *Escherichia coli* maltose binding protein (MBP) and TEM1 β-lactamase (BLA) to create β-lactamase enzymes whose catalytic activity was modulated by the presence of maltose. Here, we present a directed evolution algorithm for nonhomologous recombination and use it to evolve switches with unprecedented differences in activity between the "on" and "off" states. We also demonstrate how the phenotype conferred by these switches can be exploited to redesign switches to respond to new effectors and to create ligand-binding proteins.

Materials and Methods

MBP–BLA Library Construction Selection and Screening. Details on library construction, library statistics, and library selection/screening are provided in Tables 4–6 and Figs. 4–8, which are published as supporting information on the PNAS web site. Random domain insertion and random circular permutation of the *BLA* gene was performed essentially as described in refs. 9 and 12. Libraries were plated on LB plates containing 50 mM maltose at different concentrations of ampicillin and incubated at 37°C overnight. Based partly on the number of library members that could grow on plates containing different amounts of maltose and ampicillin, colonies from different plates were screened for a maltose dependence on nitrocefin hydrolysis activity (12). A 2-fold difference in the initial rate of nitrocefin activity was required for a clone to be examined further. For Libraries 2–5, all switches were identified from plates with 50 mM maltose and 5 μg/ml ampicillin. For Libraries 6 and 7, all switches originated from plates with 50 mM maltose and 200 μg/ml ampicillin.

Creation of Ligand-Binding Site Library in MBP317–347 (Library SB3). A library of variants of MBP317–347 was constructed in which each of the codons coding for the five positions (D14, K15, W62, E111, and W230) was randomized. Five sets of primers (in which the above codons were varied as 5′-NNK-3′) were used to amplify fragments of the *MBP317–347* gene. These fragments were assembled into a full gene by overlap extension PCR (14) in a single PCR. The assembled gene library was inserted between the BamHI and SpeI sites of pDIM-C8 to create a library of 1.58×10^7 transformants.

Selection and Screening of Library SB3. The library was plated on LB plates containing 256 μg/ml ampicillin and various amounts of

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: BLA, TEM1 β-lactamase; MBP, *E. coli* maltose binding protein; MIC, minimum inhibitory concentration.

*A patent application relating to this work has been filed for which M.O. and G.G. are coinventors.

†To whom correspondence should be addressed. E-mail: oster@jhu.edu.

sucrose. Selected colonies that grew were screened for sucrose dependence on nitrocefin hydrolysis as described above for the MBP–BLA libraries. Additional library information is provided in Table 7, which is published as supporting information on the PNAS web site.

Protein Characterization. His-tagged proteins were purified as described in ref. 12, except for protein SBP(5-7), which had a 6×His-tag and was expressed from pDIM-C8 instead of pET24b. All enzymatic assays were performed in the presence of 100 mM sodium phosphate buffer, pH 7.0. Enzyme stock was added to 2.25 ml of buffer (containing the saccharide, if desired). After incubation at the desired temperature for 5 min, 0.25 ml of 10× substrate was added and the absorbance at the appropriate wavelength was recorded by using a Cary50 UV–visual light spectrophotometer. In determining the time until half the substrate was hydrolyzed, time measurements <90 min were measured directly and time measurements >90 min were projected based on the steady-state rate. The steady-state rate did not decrease significantly over the 90 min; thus, protein instability was not an issue. Dissociation constants were determined as described in ref. 12.

The oligomeric state of MBP317–347 was determined by analysis of size exclusion chromatography data using a prepacked column of Superose 6 (Amersham Pharmacia) with a separation range of 5–5,000 kDa on a Amersham Pharmacia FPLC system. The mobile phase was phosphate buffer at pH 7.0 (0.1 M sodium phosphate/0.15 M NaCl) with or without 5 mM maltose, and the flow rate was set at 0.5 ml/min. Elution peaks were detected by UV absorbance at 254 nm. The column was calibrated with ribonuclease A (13.7 kDa), albumin (67 kDa), aldolase (158 kDa), and catalase (232 kDa) as molecular mass standards.

Results

The *malE E. coli* gene codes for MBP, a periplasmic binding protein that undergoes a large conformation change upon binding maltose (15). The *bla* gene encodes BLA, a protein that hydrolyzes β-lactam antibiotics. We have previously described three switches created by recombining the *bla* and *malE* genes. Switches EE and T164–165 were isolated from a library in which *bla* was randomly inserted into *malE* (9) (Fig. 1, Library 1). A superior switch (RG13) was isolated from a library in which random circular permutations of *bla* were randomly inserted into *malE* (12). The catalytic activity of RG13 was compromised in the absence of maltose but increased in the presence of maltose (3-fold increase in k_{cat}; 25-fold increase in k_{cat}/K_m). The library that produced RG13, however, only covered a small fraction of the possible sequence space accessible by recombination of the two genes. If one considers only the number of in-frame combinations of the two genes, the possible number of variants is 2.92×10^5 (the product of 1,110 possible locations in *malE* to insert *bla* and 263 possible circular permuted variations of *bla* that will be in-frame at each location). However, this calculation only includes "perfect" insertions and circular permutations. The method of creating the libraries generates deletions and tandem duplications at the site of insertions and the site of circular permutation (9). Even if one only considers those variants in which up to 10 codons are either tandemly duplicated or deleted, either in the circularly permuted gene or at the site of insertion, the number of possible variants increases to $>10^8$. The library that produced RG13 was estimated to contain, at most, 2.7×10^4 of these permutations. Thus, the vast majority of this sequence space was unexplored.

Iterative Library Construction and Identification of Switches. We sought to explore this unexamined sequence space in a directed fashion. The rationale behind our strategy was that sites for insertion and circular permutation found in known switches were more likely to be useful for creating switches than sites chosen at random.

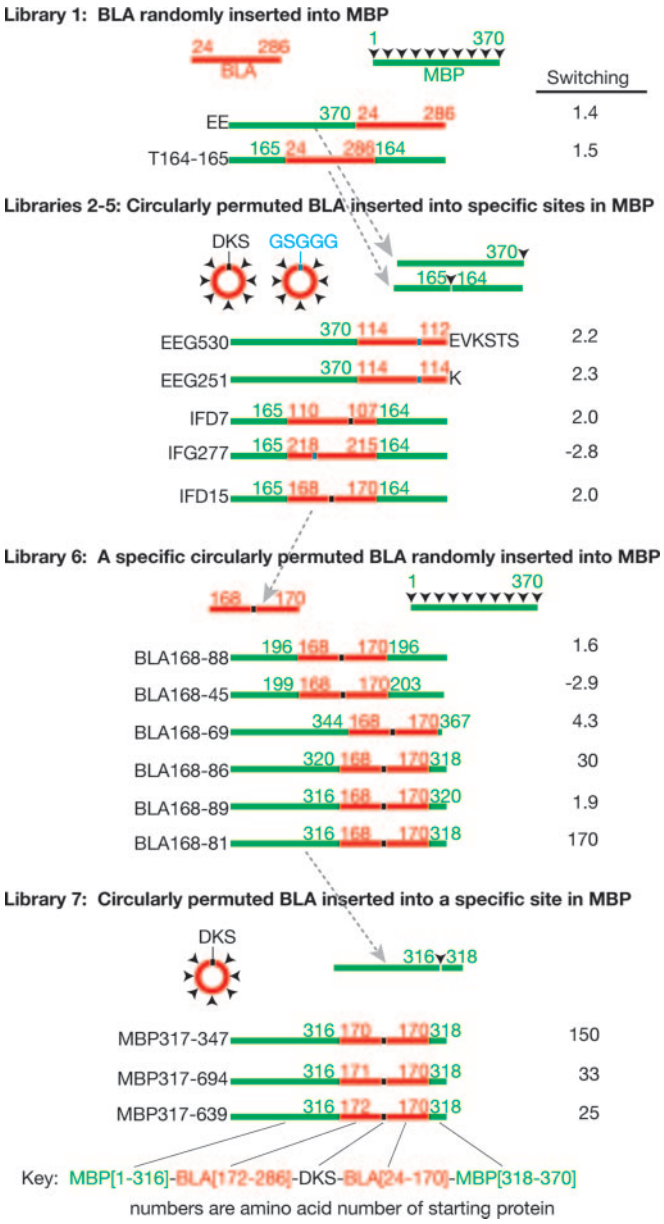

Fig. 1. Schematic depiction of library construction schemes and of the switches identified. The arrowheads indicate the sites of insertion. Multiple arrowheads on a linear gene indicate random insertion was performed. Multiple arrowheads on a circular gene indicate that random circular permutation was performed. Gray dashed arrows indicate the switch on which successive libraries are based. Switches were characterized by assaying the nitrocefin hydrolysis activity (50 μM nitrocefin; 25°C) of the soluble fraction of cell lysates in the presence and absence of maltose. Switching was quantified as a ratio of initial rates. A positive number indicates faster hydrolysis in the presence of maltose. A negative number indicates slower hydrolysis in the presence of maltose.

In our first libraries (Libraries 2–5), the sites of *bla* insertion in *malE* of EE and T164–165 were kept fixed and libraries of circular permuted *bla* genes were inserted at these locations (Fig. 1). The *bla* gene was circularly permuted such that it coded for either a GSGGG or a DKS linker between the original N- and C- termini. The DKS linker was chosen because it was previously identified as a beneficial tri-peptide linker for BLA circularly permuted at residue 216 (16). The GSGGG linker was designed to be of sufficient length to connect the termini without perturbing BLA

structure. The four libraries underwent selection for ampicillin resistance in the presence of maltose, followed by screening for maltose-dependent nitrocefin hydrolysis activity as described in ref. 12. A total of five switches were identified with improved switching activity (Fig. 1), including one in which maltose was a negative effector (IFG277). One switch (IFD15) was permuted such that residues 168–170 of BLA were tandemly duplicated. Residues 168–170 are part of the Ω-loop associated with the active site of the enzyme that includes a key catalytic residue, Glu-166 (Fig. 9*a*, which is published as supporting information on the PNAS web site). Several tolerated short deletions and insertions in the Ω-loop have been described, some of which result in altered substrate specificity (17, 18).

The fact that BLA could be permuted so near the active site without elimination of activity, combined with the notion that a connection between BLA and MBP near the active site of BLA would be more likely to produce switches with superior properties, lead us to choose this particular circular permutation of the *bla* gene for Library 6. Library 6 consisted of this circular permuted variation of *bla* randomly inserted into *malE*. From this library, several switches were identified, including BLA168–69 in which 22 residues near the C terminus of MBP were deleted. BLA168–81, whose initial rate of nitrocefin hydrolysis increased two orders of magnitude in the presence of maltose, had the circular permuted BLA inserted in place of residue 317 of MBP (Fig. 9 *b* and *c*). Interestingly, RG13 also consists of an insertion in place of residue 317 but with a different circular permutation of BLA. Thus, to explore insertions of circular permuted variants of BLA that replace residue 317 of MBP, Library 7 was constructed. Three switches with sequences very similar to BLA168–81 were identified from this library (Fig. 1).

***In Vitro* Characterization of Switches.** His tags were added to the C terminus of switches MBP317–347, MBP317–639, MBP317–694, and BLA168–81, and the proteins were purified to >95% purity by nickel-affinity chromatography. The enzymatic activity of the switches was characterized by using the colorimetric substrate nitrocefin. Only sugars known to bind MBP were effectors; sucrose, galactose, and lactose had no effect on the rate of hydrolysis. The rate of nitrocefin hydrolysis was profoundly affected by maltose (Fig. 2*a*). None of the four switches' enzymatic activity obeyed Michaelis–Menten kinetics. Catalysis was characterized by a small burst lasting on the order of several minutes followed by a slower steady-state rate (Fig. 2*b*). The size of the burst was much greater than one mol product per mol of enzyme and was consistent with a branched pathway mechanism involving substrate-induced progressive inactivation (19). Such kinetics have been observed in class A β-lactamases on substrates with bulky side-chain substituents (20) that orient toward the Ω-loop (21, 22) as well as in mutants of *Staphylococcus aureus* PC1 β-lactamase in which the Ω-loop had been perturbed (23). Thus, it is not surprising that the four switches have non-Michaelis–Menten kinetics, because the BLA domain of all four are circularly permuted (and thus perturbed) in the Ω-loop. Similar burst kinetics were seen in the presence of maltose; thus, the proposed substrate-induced inactivation cannot be an explanation for the compromised activity in the absence of maltose.

A complete kinetic characterization was not necessary to have an effective description of the maltose dependence on enzyme activity. The amount of time necessary to convert half of the substrate to product was characterized. Preliminary characterization indicated that MBP317–347 had the largest switching activity (Fig. 2*c*), and this switch was characterized in more detail. MBP317–347 exhibited a maltose dependence on activity over a wide range of protein concentrations (Fig. 2*d*). Because the catalytic activities with and without maltose differed so greatly, there was only a limited protein concentration range in which both activities could be measured. In this range, the amount of time necessary to convert half the substrate to product was 240–590 times greater in the absence of maltose than in its presence. The catalytic activity of MBP317–347 in the presence of maltose was not as great as that of wild-type BLA. For example, the amount of time necessary for 1 nM of maltose-bound MBP317–347 to hydrolyze one-half of 100 μM nitrocefin is 39-times longer than the amount of time it takes wild-type BLA to do the same. Interestingly, ampicillin and nitrocefin binding to MBP317–347 had a 10- to 20-fold negative effect on maltose affinity (Table 1). Thus, maltose-binding must have an equivalent negative

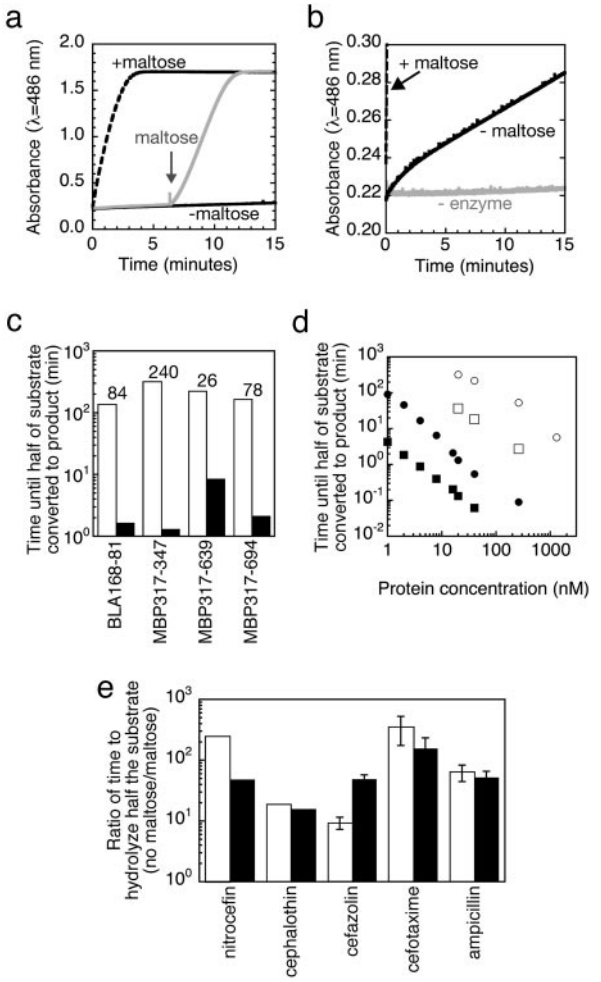

Fig. 2. Enzymatic activity of switches. (*a*) Hydrolysis of 80 μM nitrocefin by 27 nM MBP317–347 in the presence and absence of maltose at 25°C. The reaction was started by the addition of nitrocefin at time 0 to samples lacking (solid lines) or containing (dashed line) 5 mM maltose. For the reaction traced by the solid gray line, 5 mM maltose was added to the reaction at ≈6 min. (*b*) Same data as *a* with a narrower range of absorbance shown. The gray line is the background rate of nitrocefin hydrolysis in the absence of enzyme. (*c*) The time necessary for the indicated switches (20 nM) to convert half of the nitrocefin (100 μM) to product at 25°C as a function of maltose. Open bars, without maltose; filled bars, with 5 mM maltose. The number above the bars is the ratio of the times. (*d*) The time necessary for MBP317–347 to convert half of the nitrocefin to product at 25°C as a function of nitrocefin, maltose, and MBP317–347 concentration. Squares, 5 μM nitrocefin; circles, 100 μM nitrocefin; filled symbols, with maltose; open symbols, without maltose. (*e*) Ratio of time necessary for MBP317–347 to convert half of substrate to product in the absence of maltose to that in the presence of 5 mM maltose as a function of substrate and temperature. Open bars, 25°C; filled bars, 37°C. Concentrations of MBP317–347/concentration of substrate: 22.6 nM/100 μM (nitrocefin), 226 nM/150 μM (cephalothin), 113 nM/200 μM (cefazolin), 453 nM/100 μM (cefotaxime), and 113 nM/200 μM (ampicillin). Error bars for cefazolin and ampicillin indicate the standard deviation in three replicates. Error bars for cefotaxime reflect an estimated 50% error in the rate in the absence of maltose used to estimate the half-life (i.e., the rate was barely above background).

Table 1. Dissociation constants of MBP317–347 and maltose at 25°C

Substrate	Concentration	K_d for maltose
No substrate	NA	0.5 ± 0.1
Ampicillin	500 μM	10.2 ± 2.4
Nitrocefin	5 μM	1.9 ± 0.2
	100 μM	6.9 ± 0.7

The K_d in the absence of substrate was determined by change in intrinsic protein fluorescence as a function of maltose concentration (24). The apparent K_d in the presence of substrate was calculated by using change in initial rates of substrate hydrolysis as a function of maltose concentration (12). NA, not applicable.

effect on ampicillin and nitrocefin affinity [for a detailed explanation, see Guntas *et al.* (12)]. Evidently, maltose-binding must considerably enhance the rate of the catalytic steps for it to have an overall positive effect on the rate of hydrolysis. Switching occurred to different extents on all substrates tested and at 25°C and 37°C (Fig. 2*e*).

The oligomeric state of MBP317–347 at 25°C was investigated by using size exclusion chromatography. This analysis was consistent with a monomer–dimer equilibrium with a dissociation constant of ≈5 μM in the absence of maltose and ≈20 μM in the presence of maltose. The importance of this apparent dimerization (and its minor maltose dependence) for the switching activity is likely minimal: The difference in activity between with and without maltose does not have a significant dependence on protein concentration (Fig. 2*d*), and all of the protein concentrations assayed are well below the dissociation constant of the dimer.

***In Vivo* Characterization of Switches.** Switch MBP317–347 conferred on *E. coli* cells a maltose-dependent β-lactam resistance phenotype (Table 2). Initial experiments were performed on LB plates. Cells expressing MBP317–347 had a 4-fold higher minimum inhibitory concentration (MIC) for ampicillin and carbenicillin in the presence of added maltose (50 or 500 μM maltose). Other sugars, including sucrose, had no effect on the MIC for ampicillin. These maltose-dependent MIC differences were not as great as might be expected based on the *in vitro* hydrolysis rate differences. However, yeast extract at the concentration present on the plates was able to partially increase the rate of nitrocefin hydrolysis (6-fold increase) in *in vitro* assays with purified MBP317–347, presumably because of the presence of low levels of maltodextrins in the yeast extract. On media lacking yeast extract (tryptone broth), the effect of maltose on MIC for ampicillin and carbenicillin was 16-fold. The improvement in MIC difference arose from a 4-fold decrease in the MIC in the absence of added maltose.

Application of Switches to the Creation of Ligand-Binding Proteins. Switch MBP317–347 connects the presence of a ligand (maltose) to a growth/no growth phenotype when cells producing MBP317–347 are plated on β-lactam antibiotics. We sought to exploit this phenotype to create switches that respond to new effectors (Fig. 3). Our hypothesis was that if the maltose-binding site of the switch was altered such that it bound a new ligand and if binding this ligand induced a similar conformational change in the switch, then the β-lactamase activity of the switch would increase. Thus, from a library in which the maltose-binding site of the switch was randomized, one could select for those members that bound a new ligand by plating in the presence of the new ligand at concentrations of β-lactam antibiotic that normally only allow growth if maltose is present.

We tested this hypothesis by attempting to convert MBP317–347 into a switch that responded to sucrose. Maltose is a disaccharide of glucose. Sucrose is a disaccharide of glucose and fructose. Neither MBP nor MBP317–347 shows any detectable binding of sucrose ($K_d \gg 50$ mM). By inspection of the crystal structure of MBP bound to maltose (15), we identified five residues (D14, K15, W62, E111, and W230) that would be proximal to fructose if MBP were engineered to bind sucrose in the same orientation that it bound maltose. A library of variants of MBP317–347 was created in which each of the five positions was randomized. This library (Library SB3) consisted of 1.58×10^7 transformants (with a theoretical degeneracy on the protein level of 4.08×10^6). Library SB3 was plated on LB plates (the effect of yeast extract was not known at the time) at 37°C in the presence of 256 μg/ml ampicillin and increasing concentrations of sucrose. In the absence of sucrose, the frequency of library members that grew was $\approx 1 \times 10^{-5}$. We speculate that the majority of these false positives likely arose from mutations that alleviated the defect in β-lactam hydrolysis in the absence of ligand and not from mutations that improve production of protein, given that changing from LB media (4-fold MIC difference) to tryptone media (16-fold MIC difference) did not improve the frequency of false positives in the SB3 library (Table 7). The frequency of colonies on plates with 0.5 mM sucrose was not statistically different from that on plates with no sucrose. However, the frequencies of colonies growing at 5 and 50 mM sucrose were $\approx 4 \times 10^{-5}$ and $\approx 2 \times 10^{-4}$, respectively. High levels of sucrose in the growth medium can improve the production of soluble BLA through a direct effect on its folding (25); thus, some of the colonies that form only on high-sucrose plates may result from improved protein production and not sucrose-induced switching.

A total of 585 colonies from sucrose-containing plates were used to inoculate 96-well plates. Lysates of these cultures were screened (by using the 96-well nitrocefin assay) for those members that catalyzed nitrocefin hydrolysis at a faster rate in the presence of sucrose. Two library members (switches 5-7 and 6-47) originating from the 0.5 mM sucrose plate were found to respond to sucrose

Table 2. MIC of antibiotics for DH5α-E cells expressing MBP, BLA, or MBP317–347

	LB plates, μg/ml				Tryptone plates, μg/ml	
			MBP317–347		MBP317–347	
Antibiotic	MBP	BLA	− maltose	+ maltose	− maltose	+ maltose
Ampicillin	2	8,192	128	512	32	512
Carbenicillin	4	>4,096	1,024	4,096	256	4,096
Cefotaxime	0.02	0.04	0.04	0.04	—	—
Cefazolin	1	16	1	2	—	—
Cephalothin	4	64	4	8	—	—

DH5α-E cells expressing the indicated proteins were plated on the indicated media with or without 500 μM maltose and incubated 18–24 hours at 37°C. All genes were under the control of the *tac* promoter on pDIM-C8. Isopropyl β-D-thiogalactoside was not added. The MBP and BLA plates did not contain maltose. We have previously shown that maltose has no effect on the MIC for cells expressing MBP or BLA (12). —, not determined.

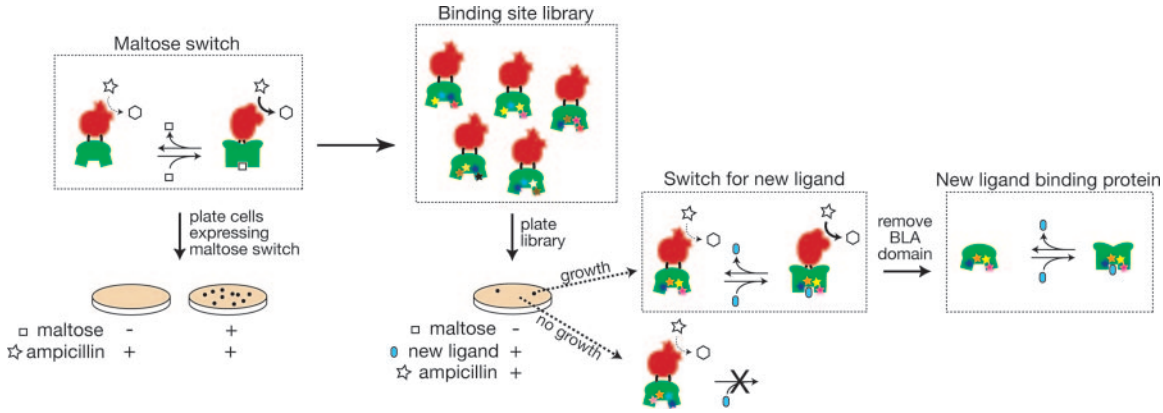

Fig. 3. Schematic depiction of the creation of switches and binding proteins from an existing switch that responds to maltose.

(Table 3). The nitrocefin hydrolysis activity of several library members originating from the 5 mM and 50 mM sucrose plates were found to respond to high levels of sucrose. These were further screened for those that responded to lower levels of sucrose, two of which were sequenced and characterized (1-59 and 1-68).

Characterization of the Sucrose Switches. The binding of sucrose and maltose to purified 5-7 and 6-47 was characterized by two different methods (Table 3). All of the switches still retained significant maltose affinity. Switch 5-7 had the highest affinity for sucrose ($K_d = 0.7\ \mu M$) and the highest specificity for sucrose over maltose (33-fold higher affinity for sucrose). No switching in response to lactose or galactose was observed. Sucrose and maltose increased β-lactamase activity by similar amounts (Table 3). However, the switching magnitude (ratio of catalytic activity in the presence of maltose to that in its absence) was less than that observed in the parental maltose switch MBP317–347. The reasons for this were examined in 5-7. In the absence of either sucrose or maltose, 5-7's nitrocefin hydrolysis activity was about 3-fold higher than MBP317–347's. The measured activity of 5-7 and MBP317–347 in the presence of bound ligand did not differ significantly. This result suggests that the nitrocefin hydrolysis activity of the apo form of 5-7 is less compromised than the apo form of MBP317–347 and that the conformation of 5-7 bound to maltose or sucrose is the same, at least as far as its effect on 5-7's β-lactamase activity. *E. coli* cells expressing 5-7 had a 2-fold higher MIC for ampicillin on tryptone plates supplemented with 500 μM sucrose.

Creation of a Sucrose-Binding Protein. The creation of ligand-binding proteins has applications in the areas of biosensors, diagnostics, chiral separations, therapeutic proteins, and basic biological studies. Periplasmic binding proteins, like MBP, have been shown to be a versatile template for the computational design of ligand binding functionalities (27, 28). We hypothesized that the mutations necessary to convert the maltose switch into one for sucrose could be used to convert MBP into a sucrose-binding protein. As such, the strategy depicted in Fig. 3 would be an alternative to more traditional methods of creating ligand-binding functionalities (e.g., phage display).

The D14L, K15F, W62Y, and E111Y mutations of 5-7 were introduced into a His-tag version of MBP to create SBP(5-7). Its affinity for maltose was the same as that of 5-7 but the affinity for sucrose was ≈10-fold less (Table 3). The conversion of MBP to SBP(5-7) is a $\gg 10^6$ improvement in binding specificity for sucrose over maltose.

Discussion

Our algorithm for creating switches involved several rounds of diversity generation followed by selective pressure. It can be viewed as a form of directed evolution, except for the technical distinction that the genetic material selected in one round is not carried forward intact into the next round. The sequence diversity in the family of MBP–BLA switches created here suggests that the sequence space accessible through random circular permutation and random insertion between two genes is rich in switching

Table 3. Sequences, ligand affinity, and switching activity of engineered proteins

						K_d for ligand at 25°C, μM					
	Amino acid number					No substrate		Nitrocefin		Switching	
Protein	14	15	62	111	230	Sucrose	Maltose	Sucrose	Maltose	Sucrose	Maltose
MBP317–347	D	K	W	E	W	nb	0.5 ± 0.1	nb	1.9 ± 0.2	1	240
5-7	L	F	Y	Y	W	0.7 ± 0.1	23 ± 13	6.7 ± 0.2	35 ± 5	82	86
6-47	L	Q	Y	Q	W	—	—	220 ± 10	3.2 ± 0.3	89	101
1-59	K	E	Y	R	W	—	—	340 ± 20	44 ± 2	28	—
1-68	L	E	Y	R	W	—	—	130 ± 24	12 ± 1	32	—
SBP(5-7)	L	F	Y	Y	W	6.6 ± 0.6	24 ± 4	n/a	n/a	n/a	n/a
MBP	D	K	W	E	W	nb	1.0*	n/a	n/a	n/a	n/a

Dissociation constants for no substrate values were determined by change in intrinsic protein fluorescence as a function of ligand concentration (24). Apparent dissociation constants in the presence of nitrocefin (5 μM) were calculated using change in initial rates of nitrocefin hydrolysis as a function of ligand concentration (12). Values for switching are ratios (without ligand to with ligand) of the time necessary to hydrolyze one half of the substrate (100 μM nitrocefin; 25°C; 20 nM protein; saturating concentration of indicated ligand). Proteins 1-59 and 1-68 were characterized in the soluble fraction of cell lysates (unknown switch concentration). nb, No binding could be detected ($K_d >>$ 50 mM). n/a, not applicable. —, not determined.
*Value is from Hall *et al.* (26).

activity, at least for these two genes. Even among the majority of variants in our libraries that do not behave as a switch, the potential for switching behavior is likely to be high. Previous studies on nonswitching fusions show that mutations that facilitate switching behavior by destabilizing the inserted protein (6) or by an unknown mechanism (5, 11, 13) can be found. Similarly, we expect that random mutagenesis or other directed evolution methodologies could be used to further improve the switches created here. Regardless of the method of diversity generation, an important hurdle in using directed evolution to evolve switches is the development of selection/screening protocols that identify conditional functionality. Although switches have been created by noncombinatorial domain insertion (4, 8, 13), the power of our combinatorial approach involving circular permutation is evident in the one- to two-order-of-magnitude-greater difference in switching activity achieved.

MBP317–347 appears to function as a monomeric allosteric enzyme. In contrast, examples of natural monomeric allosteric enzymes are very rare (29, 30), and the models that dominate our view of allostery involve interactions between subunits of an oligomer having a symmetric assembly. Such assemblies are favorable because they allow homotropic cooperative allosteric transitions that enable large changes in response over a narrow range in ligand concentration. However, our switches were designed to exhibit a heterotropic allosteric effect for which no requirement for (or advantage of) oligomeric/symmetric structures is apparent. The heterotropic allosteric effect exhibited by MBP317–347 and maltose is larger than that of many natural allosteric enzymes. The catalytic activity of *E. coli* aspartate transcarbamoylase, a classic example of an allosteric enzyme, increases just 4.5-fold in the presence of the heterotropic positive effector ATP and decreases 4-fold in the presence of heterotropic negative effector CTP (31, 32). Enzymes with allosteric effects as large as two orders of magnitude are not common, but they do exist. For example, the binding of phosphonenolpyruvate to *Bacillus stearothermophilus* phosphofructokinase decreases its affinity for its substrate by ≈100-fold (33).

The mechanism by which information is transmitted from the maltose-binding site to the catalytic site is likely to be complex and involve the coupling of the binding site and the catalytic site through a chain of propagating interactions (34–38). As expected, the switching appears to depend on a conformational change in the MBP domain of the switch. Maltose-induced fluorescence quenching and wavelength shift in our switches are much like that seen in MBP (24), suggesting that the MBP domain in MBP317–347 is undergoing a similar conformational change. The cyclic sugar β-cyclodextrin binds MBP but with a smaller conformational change than that exhibited upon maltose-binding (39). β-cyclodextrin binds to MBP317–347 (as evidenced by fluorescence wavelength shift) but does not activate MBP317–347; instead, a small negative effect on the catalytic rate is observed.

The switches described here establish a connection between a protein-ligand interaction and a selectable phenotype (antibiotic resistance), which enables (*i*) the use of genetic selections for purposes of manipulating molecular recognition between protein and ligand and (*ii*) the detection of ligand and ligand-binding by using a simple growth/no growth assay. The modular architecture of our switches facilitated the creation of a sucrose-responsive switch from an existing maltose-responsive switch. However, as the input and output domains are coupled, manipulation of the input domain can adversely affect the switching behavior, as illustrated by the decreased magnitude of switching in the sucrose switches.

The use of an engineered enzymatic switch as a ligand-binding reporter (Fig. 3) is an attractive strategy for engineering ligand-protein interactions. Chief among the advantages of this approach are the ability to use a genetic selection and that the selection occurs in the cellular environment. However, this method also has a number of limitations, including an appreciable false-positive rate and the requirement for an existing switch derived from the protein whose ligand-binding property is to be altered. In addition, it will be difficult to identify ligand-binding proteins that express much poorer than the original switch or do not switch to the higher level of activity upon binding the new ligand.

We have demonstrated how the functions of MBP and BLA (two unrelated proteins) can be substantially coupled through a directed evolution algorithm involving random domain insertion and random circular permutation. 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) (40); thus, many proteins are presumably amenable to circular permutation. Therefore, we believe our combinatorial approach has the potential to be a fairly general method for coupling many different proteins' functionalities, enabling a wide variety of applications. For example, the coupling of two proteins' functions might be used to create proteins whose function is modulated by the cellular state and thus exhibit selective therapeutic properties (e.g., artificial transcription factors, conditionally toxic proteins, proteins for altering signal transduction pathways, and proteins for targeted drug delivery). In addition, such switches represent an important addition to the synthetic biologist's toolbox for creating programmable cells for biotechnological and bioengineering applications (41) because they directly link the protein's specific activity to the cellular state.

We thank Theresa Good for use of the FPLC system. This work was supported by a grant from the National Institutes of Health.

1. Buskirk, A. R. & Liu, D. R. (2005) *Chem. Biol.* **12,** 151–161.
2. Kaern, M., Blake, W. J. & Collins, J. J. (2003) *Annu. Rev. Biomed. Eng.* **5,** 179–206.
3. Vuyisich, M. & Beal, P. A. (2002) *Chem. Biol.* **9,** 907–913.
4. Baird, G. S., Zacharias, D. A. & Tsien, R. Y. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11241–11246.
5. Doi, N. & Yanagawa, H. (1999) *FEBS Lett.* **457,** 1–4.
6. Tucker, C. L. & Fields, S. (2001) *Nat. Biotechnol.* **19,** 1042–1046.
7. Dueber, J. E., Yeh, B. J., Chak, K. & Lim, W. A. (2003) *Science* **301,** 1904–1908.
8. Radley, T. L., Markowska, A. I., Bettinger, B. T., Ha, J. H. & Loh, S. N. (2003) *J. Mol. Biol.* **332,** 529–536.
9. Guntas, G. & Ostermeier, M. (2004) *J. Mol. Biol.* **336,** 263–273.
10. Gryczynski, U. & Schleif, R. (2004) *Proteins* **57,** 9–11.
11. Buskirk, A. R., Ong, Y. C., Gartner, Z. J. & Liu, D. R. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 10505–10510.
12. Guntas, G., Mitchell, S. F. & Ostermeier, M. (2004) *Chem. Biol.* **11,** 1483–1487.
13. Skretas, G. & Wood, D. W. (2005) *Protein Sci.* **14,** 523–532.
14. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. & Pease, L. R. (1989) *Gene* **77,** 61–68.
15. Quiocho, F. A., Spurlino, J. C. & Rodseth, L. E. (1997) *Structure (London)* **5,** 997–1015.
16. Osuna, J., Pérez-Blancas, A. & Soberón, X. (2002) *Protein Eng.* **15,** 463–470.
17. Hayes, F., Hallet, B. & Cao, Y. (1997) *J. Biol. Chem.* **272,** 28833–28836.
18. Osuna, J., Yanez, J., Soberon, X. & Gaytan, P. (2004) *Nucleic Acids Res.* **32,** e136.
19. Waley, S. G. (1991) *Biochem. J.* **279,** 87–94.
20. Citri, N., Samuni, A. & Zyk, N. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 1048–1052.
21. Chen, C. C., Rahil, J., Pratt, R. F. & Herzberg, O. (1993) *J. Mol. Biol.* **234,** 165–178.
22. Strynadka, N. C., Adachi, H., Jensen, S. E., Johns, K., Sielecki, A., Betzel, C., Sutoh, K. & James, M. N. (1992) *Nature* **359,** 700–705.
23. Chen, C. C. & Herzberg, O. (1999) *Protein Eng.* **12,** 573–579.
24. Hall, J. A., Gehring, K. & Nikaido, H. (1997) *J. Biol. Chem.* **272,** 17605–17609.
25. Bowden, G. A. & Georgiou, G. (1990) *J. Biol. Chem.* **265,** 16760–16766.
26. Hall, J. A., Ganesan, A. K., Chen, J. & Nikaido, H. (1997) *J. Biol. Chem.* **272,** 17615–17622.
27. Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. (2003) *Nature* **423,** 185–190.
28. Marvin, J. S. & Hellinga, H. W. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 4955–4960.
29. Sintchak, M. D., Arjara, G., Kellogg, B. A., Stubbe, J. & Drennan, C. L. (2002) *Nat. Struct. Biol.* **9,** 293–300.
30. Ludwig, M. L. & Matthews, R. G. (2002) *Nat. Struct. Biol.* **9,** 236–238.
31. Gerhart, J. C. & Pardee, A. B. (1962) *J. Biol. Chem.* **237,** 891–896.
32. West, J. M., Tsuruta, H. & Kantrowitz, E. R. (2004) *J. Biol. Chem.* **279,** 945–951.
33. Ortigosa, A. D., Kimmel, J. L. & Reinhart, G. D. (2004) *Biochemistry* **43,** 577–586.
34. Luque, I. & Freire, E. (2000) *Proteins* **4,** Suppl., 63–71.
35. Pan, H., Lee, J. C. & Hilser, V. J. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 12020–12025.
36. Yu, E. W. & Koshland, D. E., Jr. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 9517–9520.
37. Suel, G. M., Lockless, S. W., Wall, M. A. & Ranganathan, R. (2003) *Nat. Struct. Biol.* **10,** 59–69.
38. Gunasekaran, K., Ma, B. & Nussinov, R. (2004) *Proteins* **57,** 433–443.
39. Evenas, J., Tugarinov, V., Skrynnikov, N. R., Goto, N. K., Muhandiram, R. & Kay, L. E. (2001) *J. Mol. Biol.* **309,** 961–974.
40. Krishna, M. M. & Englander, S. W. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 1053–1058.
41. Kobayashi, H., Kaern, M., Araki, M., Chung, K., Gardner, T. S., Cantor, C. R. & Collins, J. J. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 8414–8419.