Converting a Periplasmic Binding Protein into a Synthetic Biosensing Switch through Domain Insertion

Lucas F. Ribeiro,1 Vanesa Amarelle,2 Liliane F. C. Ribeiro,3 and Maria-Eugenia Guazzaroni1

1Department of Biology, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, University of São Paulo, São Paulo, SP, Brazil
2Department of Microbial Biochemistry and Genomics, Biological Research Institute Clemente Estable, Montevideo, Uruguay
3Department of Biochemistry and Immunology, Faculdade de Medicina de Ribeirão Preto, University of São Paulo, São Paulo, SP, Brazil

Correspondence should be addressed to Lucas F. Ribeiro; lucasfribeiro@usp.br

Received 29 October 2018; Accepted 17 December 2018; Published 3 January 2019

1. Introduction

A biosensor consists essentially of an input module, responsible for interacting with the target molecule, and an output module which transforms the molecule recognition into a detectable signal [1]. Over the past decade, the interest in developing biosensors capable of sensing and responding to small molecules has shown tremendous progress, fuelled by the desire to detect disease biomarkers, pathogens, and environmental toxins, to measure metabolite concentration, to create efficient high throughput screening methods, and also to generate therapeutic response triggered by a specific small molecule [2, 3]. Despite the great biotechnological potential, there is no general strategy for the construction of biosensors. Many of the current methods use a limited repertoire of naturally occurring ligand-binding proteins to couple the binding of the target molecule to the output signal, restricting the scope of target molecules that can be detected [4].

Proteins possess properties that make them ideal recognition modules, such as impressive specificity, affinity, and versatility. In protein biosensor development, two general approaches can be highlighted to convert a binding event into a detectable signal. (i) The first approach encompasses...
the immobilization of the protein on a piezoelectric, optical, electrochemical, or electrochemiluminescence device. In this case, the binding events are recorded by the difference in a physicochemical change. This approach is known as a two-component system, and it has the ability to detect molecules that cannot be imported into the cytoplasm. Nonetheless their use as biosensors is limited by the risk of cross-talk, surface adsorption, and the requirement of extra detection equipment [5]. (ii) The second approach involves a single protein that can be used as both recognition and transduction module. Compared to the two-component systems, this arrangement of sensor and effector in one molecule is simpler and more effective and it reduces potential issues associated with surface adsorption and the dependency of the complex and expensive detection equipment [6]. This strategy is ideal for whole-cell biosensor applications [7]. Whole-cell biosensors can provide the advantages of rapid and sensitive analysis for in situ monitoring with cells [8–10]. Single-protein biosensors can be expanded through the engineering of proteins in which the molecular recognition is coupled with a detectable protein function.

A promising approach to design new generation single-protein biosensors is to expand the toolkit of the allosteric molecules known as “protein switches.” A typical protein switch is a biomolecule that can change between two or more distinct conformations (or conformational ensembles) in response to a specific stimulus [11, 12]. These changes modulate their active state – output – (e.g., enzyme activity, ligand affinity, fluorescence, and oligomeric state) in response to a binding event or physical signal – input – (e.g., small molecule, pH, covalent modification, and light). A usual approach to design switchable proteins has been to fuse a protein able to recognize an input signal (e.g., periplasmic binding proteins) with a protein whose function one desires to create an input-dependent response (Figures 1(a)–1(c)). Since protein switches can transduce an input signal into a functional response, they are logical targets to build biosensors.

Biological systems can be described as an interacting network of molecules organized in complex circuits. Protein switches are key components able to couple cellular functions. Their behavior is similar to the natural allosteric proteins, exhibiting remarkable attributes that make them an extraordinary model to design biosensors, such as high specificity and affinity, reversible signal transduction, versatility, and fast response, acting in millisecond to microsecond timescale [13, 14] which is faster than inducible gene expression-based systems (seconds to hours) [15]. In addition, unlike a single domain with linear response, allosteric switches can produce a cooperative connection of the input, leading to a finer adjustment of the output. Thus, components with highly sensitive switch behavior resemble a digital response, providing an input detection threshold in which small changes in the input concentration lead to large changes in the output response (Figure 1(e)). All these attributes are important to continuous and real-time molecule detection even inside living cells.

Synthetic proteins switches are engineered to show a user-defined input and output recognition/response. Engineered protein switches have been used in a wide variety of applications such as biosensors [6, 16–19], cancer or diabetes therapeutics [20, 21], biomass degradation [22, 23], recognition of cell signaling elements [24–26], and control of gene expression and genome editing [27–30]. However, the main challenge to design a protein switch is overcoming the problem of how to couple input and output functions, both physically and functionally, so that binding of the analyte produces a detectable signal. The science to engineer protein switches by coupling any desired input to output domains would enable the rewiring of cellular circuitry according to bioengineer’s goals.

Despite the wide potential, protein switches have not been extensively explored because of the scarcity of universal engineering strategies and the difficulty to design a protein that responds to a signal unrelated to its function, becoming more (or less) active in the presence of this signal. In this review we will discuss recent studies that used protein engineering approaches, domain insertion and directed evolution, to recombine nonhomologous proteins generating molecular switches. We will be emphasizing those studies that expanded the toolkit for allosteric switches using the superfamily of proteins known as periplasmic binding proteins (PBPs), whose members are able to bind to diverse ligands. A summary of the engineered switches, along with their properties, is provided in Table 1.

2. Engineering Switches by Protein Domain Insertion

Protein domains are evolutionarily conserved polypeptide units that usually present independent functional or structural properties. More than two-thirds of the proteins found in prokaryotes and eukaryotes contain multiple domains [31]. Moreover, protein domains can act as structural reservoirs to generate new protein architectures [32] and can be used as building blocks to design new proteins with expanded biotechnological applications [11, 12, 22, 23, 33–37]. Multidomain proteins can be engineered by domain insertion in such a way that a domain (insert) is spliced into another domain (acceptor) either at a specific position or by random insertion. Structural coupling among the combined domains can emerge from the recombination of insert and acceptor domains, with the emergence of new functions [38]. The first example of successful insertion of one protein into another was published in 1990, when Ehrmann and colleagues inserted the alkaline phosphatase from Escherichia coli into the membrane protein MalF also from E. coli. This engineered chimera was constructed as a tool for examining membrane protein topology [39]. Furthermore, different from end-to-end fusion in which domains are linked by a single contact point, when two domains are fused by domain insertion, they are linked by peptidyl bonds at two contact points with more intimate connection between the two proteins. This double contact can increase the protein stability of generated chimeras [23, 35, 40, 41], and the intimate connection can be used to couple the functions of two fused proteins.

Proteins switches can be designed to emulate the behavior of natural allosteric proteins, but with user-defined input and
output signals. Allosteric protein switches can be created by fusing two domains in such a way that the activity of the output domain is regulated by the input domain’s recognition of an input signal [12, 34] (Figures 1(a)-1(c)). A significant challenge for designing novel switches is finding the correct match between the input/output domains which allows the signal/response coupling (Figure 1(d)).

For the creation of protein switches by domain insertion, the acceptor protein has to be a discontinuous protein, its linear sequence can be interrupted by another inserted protein, which is fairly prevalent in nature [34, 42]. In addition, the N- and C-terminals of the inserted protein should be close enough to permit the insertion without disrupting the protein structure [34]. This condition is not as restrictive since at least 50% of all single domain proteins have their N- and C- termini proximal (<10 Å) [11, 34, 43]. However, the exact fusion to create a switch is difficult to predict. Thus, a directed evolution approach is preferred in which an insert gene is randomly inserted into an acceptor gene, to create a library of gene fusions encoding fused proteins. These fused proteins are then subjected to selection or screens to find chimeric proteins that possess signal-dependent response [11].

Random insertion libraries are usually created by DNase I or S1 nuclease [44], multiplex inverse PCR (MIP) [11], or in vitro transposition [45] (Figure 2). Nuclease-digestion generates a single double-stranded break throughout the plasmid that encodes the acceptor gene, while MIP uses designed primers to “open up” the plasmid to facilitate domain insertion. MIP can be used for random domain insertion (i.e., inverse PCR at each codon) or for rational design (i.e., achieved by inverse PCR only at designated positions). In in vitro transposition an engineered transposon is randomly inserted into the plasmid carrying the acceptor gene by an in vitro transposase reaction. Then, the transposon is removed by restriction digestion and single break plasmids are size selected. Both nuclease-digestion
| Function modulated in vivo | Input domain (PBPs) | Output domain | Method of creating switch | Input signal | Kd (μM) | Variant name: switch effect | Ref |
|----------------------------|---------------------|---------------|---------------------------|-------------|--------|-------------------------------|-----|
|                            | MBP                 | BLA           | Random insertion of BLA into MBP | Maltose     | 3.2    | T164-165: 1.6-fold$^c$         | [54] |
|                            | MBP                 | BLA           | Circular permutation of BLA and random insertion into MBP | Maltose     | 1.7    | T164-165-H: 1.8-fold$^c$       | [55] |
|                            | MBP                 | BLA           | Iterative circular permutation and random insertion of BLA into MBP | Maltose     | 0.5    | RG13: 25-fold$^c$              |     |
|                            | MBP                 | BLA           | Five residues in the maltose binding site were randomized generating a switch variant that binds to sucrose | Maltose     | 0.7    | MBP317–347: 82-fold$^c$        |     |
|                            | GBP                 | BLA           | cpBLA was randomly inserted into the PBP | Sucrose     | ND     | MRD2 cd9: 2-fold$^c$           |     |
|                            | RBP                 | BLA           | Random insertion of BLA into XBP and linker optimization | Glucose     | ND     | Pl. F10: 72-fold$^c$           | [33] |
|                            | XBP                 | BLA           | Site-directed mutagenesis of I329 residue of RG13 | Ribose      | 0.67   | I329A: 23-fold$^d$             |     |
|                            |                     |               |                           | Xylose       | 0.63   | I329K: 32-fold$^d$             |     |
|                            |                     |               |                           |             | 0.55   | I329W: 20-fold$^d$             |     |
|                            | MBP                 | BLA           | Disulfide bonds were rationally introduced in RG13 | Maltose (+GSH) | ND     | RG13-AND2: 2-fold$^d$          | [58] |
|                            | MBP                 | BLA           | Linker modification of nonallosteric MBP-BLA (c4) to search for emergence of allostery through modulation of the conformational entropy | Maltose     | ND     | RG13-ORN2: 8-fold$^d$          |     |
|                            | MBP                 | BLA           | Point mutations at selected residues of the maltose binding pocket of MBP317-347 | Maltose (heat or OH$^-$) | ND     | RG13-AND2: 3.5-fold$^d$        | [59] |
|                            | MBP                 | BLA           | Linker modification of nonallosteric MBP-BLA (c4) to search for emergence of allostery through modulation of the conformational entropy | Maltose     | ND     | c4-4G: 32-fold$^d$             | [60] |
|                            | MBP                 | BLA           | Point mutations at selected residues of the maltose binding pocket of MBP317-347 | Maltose     | 499    | E153D: 2,360-fold$^f$          | [61] |
| Function modulated in vivo | Input domain (PBPs) | Output domain | Method of creating switch | Input signal | Kd (µM) | Variant name: switch effect | Ref |
|---------------------------|---------------------|---------------|---------------------------|-------------|--------|---------------------------|-----|
| **MBP**                   | GFP                 | Mal-β:2.8-fold | Malose                   | 2.8         |        |                           |     |
| **TMBP**                  | GFP                 | Malose:0.053  | Tre-C04:6.3-fold          | 0.053       |        |                           |     |
| **PhnD**                  | GFP                 | 2AEP          | EcPhnD90- cpGFP, L1ADΔΔ, L297R/L301R:1.5-fold | 37          |        |                           |     |
| **MBP**                   | GFP                 | 4.5           | MBP165-cpGFP:0.2-fold    | 4.5         |        |                           |     |
| **MBP**                   | GFP                 | 5             | MBP165-cpGFP, PPFFY:2.5-fold | 3           |        |                           |     |
| **MBP**                   | GFP                 | 1.3           | MBP175-cpGFP, L1-HL:0.5-fold | 1.3         |        |                           |     |
| **Fluorescence**          | MglB                | 600           | FLII12 P-Glu-600 μ: 2.66-fold | 600         |        |                           |     |
| **GltI**                  | YFP                 | Glutamate:1   | FLII18P-E1:μ: 3.8-fold | 1           |        |                           |     |
| **QBP**                   | CFP                 | Arginine:2100 | QBP/Citrine/CFP:1.3-fold | 2100        |        |                           |     |
| **GltI**                  | GFP                 | Ornithine:2000 | QBP/Citrine/ECFP:1.1-fold | 2000        |        |                           |     |
| **GltI**                  | GFP                 | Glutamate:107 | Glt253.L1L/V/L2NP:4.5-fold | 107         |        |                           |     |
| **GltI**                  | RFP                 | Aspartate:145 | Glt253.L1L/V/L2NP:2.5-fold | 145         |        |                           |     |
| **Enzymatic activity**    | XBP                 | Glucose:600   | FLII24 P-Glu-600μ: 2.66-fold | 600         |        |                           |     |
| **XBP**                   | XynA                | XynA:1.6      | 2621B: 1.49-fold          | 1.6         |        |                           |     |
| **Control of gene expression** | MBP                 | XynA:1.6      | XynA-2621B: 1.49-fold     | 1.6         |        |                           |     |

* Studies in which many variants were constructed, and only those displaying the best switch effect were considered.

**MBP, maltose binding protein; GBP, glucose binding protein; RBP, ribose binding protein; XBP, xylose binding protein; TMBP, trehalose/maltose-binding protein; PhnD, phosphonate-binding protein; MglB, glucose/galactose-binding protein; QBP, glutamine binding protein; GltI, glutamate-binding protein.

277A:2-fold | 270A:2-fold | 335P:2.4-fold

* in vitro or **in vivo assay values in the presence of the effector divided by the value in the absence of the effector.
and in vitro transposition methods can generate insertions outside of the coding sequence, deletions, and a significant fraction of the library harboring out-of-frame and undesired fusions, even though small in-frame deletions contribute to increasing library diversity. Therefore, both strategies are best combined with a robust screening system. The inserted gene, encoding for the input or output domain, is generally blunt ligated into the open plasmid generating the random domain insertion library. In addition, circular permutation of the insert gene can be used to create another layer of diversity, thus increasing the overall diversity of the protein switch library [11, 46].

3. Periplasmic Binding Proteins as Input Domain Provide Immense Ligand-Binding Diversity

Periplasmic binding proteins are a superfamily of bacterial receptors that mediate chemotaxis and solute uptake [47, 48]. PBP s are fundamental components of ATP-binding cassette (ABC) transport systems and are located in the periplasm of Gram-negative bacteria. In Gram-positive bacteria and archaea, the periplasmic binding component is replaced by a membrane bound lipoprotein, which acts in a similar manner. PBPs have been identified for a wide variety of ligands including amino acids, dipeptides, oligopeptides, carbohydrates, lipids, peptidoglycans, vitamins, drugs, metabolic products, quorum sensing molecules, hormones, metal ions, and anions [49, 50]. PBP-structure has also been identified in domains of DNA repressors such as LacI, eukaryotic receptors such as GluR2, and enzymes [47, 48]. Members of the PBPs superfamily show high sequence diversity, but protein structure is usually conserved [49]. They consist of two domains linked by a flexible hinge region being the binding site located at the interface between the two domains. PBPs may adopt two different conformations (Figure 3): an open conformation when no ligand is bound (apo-form) and a closed conformation upon binding (holo-form). These conformations are interconvertible by a relatively large bending movement around the hinge region [51]. Two different structural classes have been described, classes I and II, which differ in their topology [49]. Currently, there are 302 PBP structures deposited in the Protein Data Bank (research made with the term “periplasmic binding protein”).

Bioengineers have been using the intrinsic properties of PBPs to build biosensors and other elements allosterically modulated [49]. Among these properties, it is possible to highlight (i) high ligand affinity (range from nM to μM), solubility, stability [52], and (ii) the large ligand-mediated conformational changes [53], which can allow molecular communication between linked domains through conformational coupling. Although conformational changes in proteins due to ligand binding are ubiquitous and integral in biological systems, it is often challenging to link ligand-induced conformational changes to a resulting biological function.

An evident limitation to engineer novel biosensors is finding a specific PBP carrying all required properties for the tailored sensor. While still difficult to build new PBPs through de novo protein design, three distinct approaches can be applied for overcoming this issue: (i) mining novel natural PBPs using genome or metagenome strategies [70, 71]; (ii) using computational design methods upon known protein
scaffolds to create new protein–ligand interfaces [72]; and (iii) the combination of computational protein design and directed evolution to generate affinity for other ligands in a specific PBP [56, 73]. Several PBPs have been created using these approaches expanding the possibilities to design novel biosensors.

4. Protein Switches Based on Antibiotic Resistance

A cornerstone of protein engineering is the development of high throughput screening/selection systems (HTS). Construction of such systems has been used not only to generate improved proteins but also to design whole-cell biosensors that translate cellular signals into quantifiable outputs [74]. One such example is the use of antibiotic resistance as output signal in which the modulation of enzymatic activity results in the alteration of cellular survival capacity in the presence of the antibiotic. This allows the screening of large libraries (e.g., > 10^7 members), performing even better than contemporary automated platforms. Therefore, a protein switch carrying an enzyme that provides resistance to antibiotic in its output domain is an attractive biosensor for both in vivo (e.g., whole-cell biosensors for HTS) and in vitro (e.g., colorimetric enzymatic assay) applications. Ostermeier and coworkers have developed a family of growth/no-growth biosensors for maltose using the TEM1 β-lactamase (BLA) enzyme as output domain. Initially, it was demonstrated that BLA could be randomly inserted into the maltose-binding protein (MBP) from E. coli to generate switch proteins. After HTS steps, a chimeric protein was selected in which BLA activity was positively modulated by maltodextrins, either in vivo or in vitro. However, this selected switch showed low dynamic modulation (≤ 1.8-fold) [54]. Next, an improved MBP-BLA was developed by random circular permutation (cp) of the BLA gene followed by random insertion of the cpBLA library into the MBP gene (Figure 4(a)) [55]. This approach generated a switch protein (termed RGI3) with BLA catalytic efficiency increased 25-fold in the presence of maltose. BLA activity in RGI3 also gained responsiveness to a second input, the Zn^{2+} ion, which worked as a negative modulator, showing that novel properties can emerge from chimerogenesis [75]. Afterwards, an iterative approach involving alternating random domain insertion and random circular permutation generated a switch with 600-fold increase in BLA activity upon maltose binding [56]. This protein switch, termed MBP317–347, has a cpBLA (BLA170) inserted into a loop of MBP containing a key active site amino acid. Furthermore, mutations were introduced into the ligand-binding site of the MBP domain of the MBP317–347, conferring the ability to bind sucrose without eliminating maltose binding [56]. In addition, MBP317–347 also showed low affinity for glucose and the potential to be used as a biosensor of metabolic biomarkers from tumors [76].

In order to explore the structural basis of local amino acid residues important to the allosterism generated by domain insertion, Choi and coworkers [61] used the engineered protein MBP317-347 for the construction of 285 mutants. MBP317-347 binding pocket was the target for all possible amino acid replacements at 15 positions, which were selected due to their known involvement with maltose binding. Interestingly, the allosteric function and the “phenotype switch” were resistant to mutations, but a few of them affected the affinity of MBP317-347 for maltose and also the switching behavior. Mutations in positions E153 and R66 caused a decrease in BLA activity in the absence of maltose, enhancing the switch effect in the presence of the effector. Results obtained in this study suggest that the design of improved switches could be achieved by focusing on the effector-binding site to generate mutations which could enhance the modulating effect.
In addition, switches sensing ribose, glucose, and xylose have also been created through fusion of BLA into ribose binding protein (RBP) (≤7-fold), glucose binding protein (GBP) (≤2-fold), and xylose binding protein (XBP) (≤4-fold) from *E. coli*, respectively [33]. Although these switches share paralogous PBPs as input domain, and some overlap in sites for switch insertion, successful switches at these sites required different circular permutations of the BLA and different linkers between the domains.

### 4.1. Multi-Input Protein Switches

In seeking to make biology “programmable,” synthetic biology approaches have been applied to engineer cells that respond to molecular signals with a tailored behavior. Many of these approaches focus on the creation of multi-input logic gates to improve the cell decision-making efficiency. Proteins have the potential to perform all 16 possible two-input logic gate operations [77]. In addition, they have proven to be an important alternative to classical transcriptional networks for logic gates design,
offering the advantages of simplicity, speed, and reduced consumption of cellular resources [25, 58].

The chimeric MBP-BLA protein RG13 showed an allosteric mechanism dependent on conformational changes in the MBP domain [55, 75]. As a consequence of RG13 chimerogenesis, the BLA domain gained a novel noncompetitive input (Zn\(^{2+}\)), and the switch effect was modulated by two-input signals [75]. In a recent work, RG13 was rationally redesigned to demonstrate the potential for a single protein to show a multi-input control and different logic gates behavior [58]. In this study a redox-control was integrated in RG13 through introducing disulfide bonds to keep the MBP domain in either an open or closed conformation (Figure 4(b)). In the closed conformation BLA activity is “on” and it is switched to “off” state when MBP changes to an open conformation. Thus, besides the maltose-regulated allostery of RG13, the switch also became dependent on the disulfide bond, which is controlled by redox agents. These RG13 variants were also applied to create AND, ORN, and YES logic gates behavior, offering a protein-based alternative to produce a functional output from multiple inputs [58]. Moreover, since an electrochemical signal can also be used to reduce disulfide bonds, the engineered RG13 variants showed a voltage-dependent switch effect [59]. These studies suggest that disulfide containing protein switches are a useful platform to create multi-input switches with synthetic biology applications, redox-controlled biosensors, and also bioelectronic sensors.

4.2. ProteinSwitchesBasedonanEnsembleModelofAllostery.

In the last decades, the classic view of allostery has expanded, and the ensemble model emerged. This model statistically describes the existence of dynamic conformational ensembles rather than considering proteins allosterically regulated as a static two-state model (active and less active). [78–80]. A protein switch can show a behavior compatible to ensemble allostery, in which the remodeling of the landscape energy is responsible for tuning the protein activity (Figure 4(d)). In order to design switchable enzymes based on an ensemble allostereic model, a nonallosteric enzyme was converted into an allosteric protein through modulation of the conformational entropy [60]. Based on the previously identified nonallosteric bifunctional chimera (termed c4 [81]) where cpBLA was inserted into MBP, the linker region was engineered to increase the conformational flexibility of the chimera in the absence of maltose (effector). These engineered linkers created intrinsically disordered regions that reduced enzyme activity in the absence of the effector. The addition of maltose stabilized active states of the enzyme, specifically at higher pH or temperature, generating a switch effect. Thus, the modulation of the conformational entropy allowed converting a nonallosteric protein into a multi-input switch which requires a particular environmental condition (temperature or pH) and the presence of the effector for its activation [60].

5. ProteinSwitchesBasedonFluorescence

Fluorescent protein biosensors (FPBs) are probes containing a sensing domain that recognizes a target molecule and a reporter module that generates the fluorescence signal [82]. These biosensors permit quantifying analytes within and between cells, both in vivo and in vitro, through changes in fluorescence properties (e.g., excitation or emission wavelengths, intensity, and lifetime of the excited state). In addition, FPBs allow the measurement of temporal and spatial dynamics of targeted cellular events in real time [82].

Single FPBs (SFPBs) are fluorescent sensors encoded by a single polypeptide chain composed of a circularly permuted fluorescent protein (cpFP) inserted into a specific ligand-binding domain [82]. Marvin and coworkers [63, 64, 66] developed SFPBs with high signal-to-noise ratio by the insertion of cpFPs into PBPs. In order to generate biosensors for disaccharides, organophosphorus compounds, and glutamate, cpGFP was inserted into E. coli maltose-binding protein MalE [64], phosphonate-binding protein PhnD [63], and glutamate-binding protein GltI [66], respectively. Residues predicted to undergo dramatic rearrangement upon ligand binding were chosen for the cpGFP insertion, and the linkers between cpGFP and PBP were optimized. Binding of ligand induces the PBP conformational changes switching the biosensor to a bright state. The GltI/cpGFP chimera, termed iGluSnFR, showed specific response for in situ glutamate signal in neurological systems from worms, zebrafish, and mice [66].

In a recent study, Wu and coworkers [67] also developed a high-affinity biosensor for glutamate termed R-iGluSnFR0.1. This protein was created replacing the cpGFP domain of iGluSnFR by a circularly permuted red FP (cRFP). Then, the switch effect was improved through linker optimization and iterative rounds of directed evolution. In addition, a second variant was constructed by circular permutation of GltI instead of RFP. The cpGltI was inserted into nonpermuted RFP at the same position as previously identified for iGluSnFR, leading also to a high-affinity biosensor for glutamate (Rexp-iGluSnFRI). Both variants, R-iGluSnFR0.1 and Rexp-iGluSnFRI, worked successfully as biosensors for glutamate when targeted to the surface of HEK-293 cells [67]. These results show that the changing of protein topology may offer a new diversity layer for engineering fluorescent biosensors.

In an elegant approach, Nadler and coworkers [62] describe a general approach, termed domain insertion profiling with DNA sequencing (DIP-seq), to identify functional metabolite responsive SFPBs. As proof of concept to validate the methodology, they created domain insertion libraries between E. coli MBP and cpGFP, since the SFPB composed by MBP and cpGFP is a benchmark to generate allosteric fusions [54, 64]. Using in vitro transposition, MBP tolerance to random cpGFP insertion was evaluated by functional screening of the library with fluorescence-activated cell sorting (FACS) in the presence of maltose. Deep sequencing analysis showed that ~57% of MBP residues contained productive insertions of cpGFP. With this approach, maltose biosensors with a greater switching dynamic than the one previously published by Marvin and coworkers [64] were found. Moreover, with the aim of developing novel biosensor for trehalose, cpGFP was inserted into the D-trehalose/D-maltose-binding protein (TMBP) from Thermococcus litoralis.
6. Protein Switches Based on Allosteric Transcriptional Regulators

Recognition of an external stimulus (input) resulting in a particular cellular response (output) is an efficient approach to developing engineered cells with biosensor functions [86]. In most organisms, the ability to respond to an input is often closely linked to the fine control of gene expression. A key strategy for optimizing this control is the design and rearrangement of regulatory elements for tuning gene expression. Although many tools have been developed, most are focused on elements that act in cis (e.g., promoters, ribosome binding sites, and terminators) and are limited by compatibility, ease of implementation, difficulty to optimize gene expression under multiple conditions, and undesirable effects in the native physiology of the microorganism [87]. Many of these problems could be overcome with the development of trans-acting DNA-binding transcription factors (TFs) that could control the expression of a target gene without changing its native regulation.

 Naturally occurring inducible TFs have been broadly used as important elements to control gene expression in response to a target molecule. However, these natural proteins have limited recognition spectrum. One approach to engineer new conditional TFs is to use directed evolution or rational design to fuse, by domain insertion, PBPs with DNA-binding domains (DBDs), leading to versions of DBDs that respond to new ligands (Figure 4(f)).

An attractive DBD family comprises the zinc finger proteins (ZFPs), especially Cys2-His2 class, since each “finger” of the ZFP binds to a distinct 3 bp DNA sequence, and multiple “fingers” can be fused in tandem to engineer new ZFPs [88]. In order to expand the toolkit of inducible TFs, Younger and coworkers [68] inserted a ZFP (BCR-ABL1), into MBP. BCR-ABL1 binds to its cognate 9 bp DNA-binding site with three tandem fingers [89] and is optimized to work as a transcriptional repressor in E. coli by sterically blocking RNA polymerase binding to the promoter. For this purpose, BCR-ABL1 was inserted after residue 316R into MBP, a position previously identified to generate an MBP/BLA switch protein [55]. The insertion generated an MBP/BCR-ABL1 inducible repressor being the expression of its target gene activated by the presence of maltose with moderate dynamic modulation (≤ 4-fold induction). Even though the MBP/BCR-ABL1 mechanism of action is not clear, the authors hypothesized that maltose binding to the MBP domain changes the ZFP domain conformation to a state not capable of binding to DNA [68]. In a different approach, directed evolution was used to randomly insert BCR-ABL1 into MBP using in vitro transposition [69]. After four rounds of screening with FACS, in the absence or presence of maltose, three unique maltose-responsive biosensors were identified (MBP insertion points: 270A, 277A, and 335P). However, even after linker optimizations all the variants showed a moderate dynamic modulation (≤ 4-fold induction) [69].

7. Protein Switches Based on Phenotype

As mentioned before, when developing protein switches by random domain insertion, the identification of responsive chimeras relies on an efficient screening method. When this screening is made based on a cellular phenotype, the switching behavior selected could be due to mechanisms other than allostery, for instance, by accumulation of the active state into the cell. In 2011, Ostermeier’s group published a study in which a library of cpBLA was inserted in residue 317 of MBP [81]. The library was screened for chimeras that confer antibiotic resistance to E. coli in the presence of maltose. Among the 34 colonies selected, only 4 of them presented an allosteric behavior. The other colonies were later called “phenotypic switches”. In this example, a higher cellular accumulation of the fused protein was detected in the presence of the inducer, indicating that this phenotype was a result of the fused protein interaction with maltose [81]. Then, they sought to investigate the mechanism of these
phenotypic switches [90] using two previously constructed chimeric variants [81], in which BLA was inserted after residue 316 of MBP with a linker (c4) or without a linker (Ph7) between BLA and residue 319 of MBP. Ph7 behaved as a phenotypic switch and c4 showed no switching behavior, suggesting that the linker might be the cause of the different performance. Therefore, they tested a set of chimeric proteins varying the linker length between BLA and MBP. Initially, two E. coli strains were transformed with the variants and there was a significant difference of the phenotypic switch effect between the strains, suggesting that the behavior was strain dependent and that cellular factors might play an important role. The authors proved that Ph7 was more susceptible than c4 to proteolysis and that in the presence of maltose both proteins were less hydrolyzed [90]. Therefore, they suggest that maltose binding to MBP partially protects proteins from being digested. Another reason for protein accumulation in the cell could be a disturbance on thermodynamic stability. The authors suggest that the longer the linker, the higher the thermodynamic stabilization in the presence of maltose. Thus, maltose binding to MBP in the fused protein seems to be responsible for increasing protein accumulation due to lower protease susceptibility and a higher thermodynamic stability.

8. Mutually Exclusive Folding

Differently from the previous examples in which the inserted protein has its N- and C-termini close, in mutually exclusive strategy the terminals are distant from each other. This configuration forbids both domains from being simultaneously folded. Thus, the free energy of one domain is used to mechanically unfold the other, creating a structural tug-of-war between the domains (Figure 4(c)) [91–94]. Linkers can be introduced between domains; however, they cannot be too long to cause uncoupling of the structure, neither too short to prevent intramolecular folding [92]. Loh's group [95, 96] applied mutually exclusive folding principles to develop protein switches induced by a mechanism known as domain swapping. Domain swapping is a structural phenomenon by which identical proteins exchange segments in reciprocal fashion to yield oligomerization. Ubiquitin (Ub) or FK506 binding protein (FKBP) was inserted into functionally inactivated point mutants of ribose binding protein (RBP) from Thermoanaerobacter tengcongensis. Ub and FKBP worked as “lever” proteins that split RBP forcing it to first unfold then refold via swapped dimer. The mutant RBP alone cannot bind ribose; only by swapping and dimerization its function is restored. The drug FK506 worked as input signal to trigger swapping activation, and RBP worked here as an output domain [95, 96].

9. Future Perspectives

The ability of coupling an input event to an output signal in a protein switch has tremendous utility for the development of biosensors and also regulation of biological systems. The approaches outlined here show how the power of domain insertion can be harnessed for fine-tuning a protein activity and to evolve allostery. However, the scourge for designing switchable proteins by domain insertion is its low-throughput nature due to the difficulties in predicting the right insertion sites for domain coupling. Therefore, through many successful examples described above it was necessary to try many different fusions, followed by extra steps of circular permutation, linker optimization, using tedious and time-consuming screenings. Thus, overcoming this limitation is paramount to expanding protein switches applications. Engineering allosteric switches through rational insertion would be an ideal approach for biosensor designing. Ranganathan's group has used a statistical coupling analysis (SCA) to identify physically connected networks of coevolving amino acids which link the functional site of a protein with a distant surface site [97, 98]. Reynolds and coworkers [98] have used this computational tool for scanning domain insertion positions in the enzyme dihydrofolate reductase (DHFR) and inserted a light-sensitive domain (LOV2) into these “hotspots” positions, which allowed the identification of a chimeric protein regulated in vivo by light. Thus, linking of existing allosteric networks in the individual protein domains could lead to allosteric connections [99] (Figure 4(g)). The large data set generated by previous work on building protein switches could serve to expand this approach.

In addition, recent studies have explored deep sequencing analysis to address how domain insertion affects allostery [27, 62]. These datasets can help us understand better the fitness landscape for protein switches and increase the success rate for domain coupling.

PBPs conformational change has shown a powerful feature for allosteric coupling in protein switches. Nonetheless, the creation of a novel biosensor depends on the availability of a PBP to bind to the target ligand, undergo conformational changes, and be able to either accept domain insertions or have an appropriate distance between the N- and C-termini to be inserted. Thus, the limited repertoire of natural PBPs available to date is a barrier for the development of switches. Several approaches can be applied to expand the number of PBPs available, such as metagenomic [70, 71], directed evolution [56, 73], and computational design strategies [72].

In the future, hence, with continuous computational advances, it may be possible to design allosterically regulated biosensors by engineering novel PBPs, and predicting insertion sites for perfect domain coupling. This suggests that there is still plenty of scope for the engineering of switch proteins. One example is the small number of engineered TFs. Combining PBPs with DNA-binding proteins (e.g., TALE or CRISPR/Cas9) to design novel inducible TFs would have extraordinary importance. In addition, many of the studies described here are specific to bacteria. The adaptability of these tools to other platforms such as mammalian cells, fungus, and plant cells would greatly increase its biotechnological impact.

Given the rapid development of switches engineered by domain insertion, it is likely that in the future a wide range of tailored biosensors, working in vivo or in vitro, will be available for a wide variety of applications.
Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by Young Research Award by the Sao Paulo State Foundation (FAPESP, award number 2015/04309-1). Lucas F. Ribeiro and Liliane F. C. Ribeiro are beneficiaries of FAPESP fellowships (award numbers 2016/18827-7 and 2016/20358-5, respectively). We sincerely apologize to authors whose work could not be included in this manuscript due to the space limitation.

References

[1] J. Ha and S. N. Loh, “Construction of Allosteric Protein Switches by Alternate Frame Folding and Intermolecular Fragment Exchange,” in Synthetic Protein Switches, vol. 1596 of Methods in Molecular Biology, pp. 27–41, Springer New York, New York, NY, 2017.

[2] A. P. F. Turner, “Biosensors: sense and sensibility,” Chemical Society Reviews, vol. 42, no. 8, pp. 3184–3196, 2013.

[3] Y. Liu, Y. Liu, and M. Wang, “Design, Optimization and Application of Small Molecule Biosensor in Metabolic Engineering,” Frontiers in Microbiology, vol. 8, 2017.

[4] J. Feng, B. W. Jester, C. E. Tinberg et al., “A general strategy to construct small molecule biosensors in eukaryotes,” eLife, vol. 4, 2015.

[5] A. I. Podgornaia and M. T. Laub, “Determinants of specificity in two-component signal transduction,” Current Opinion in Microbiology, vol. 16, no. 2, pp. 156–162, 2013.

[6] B. C. Stanton, V. Siciliano, A. Ghodasara et al., “Systematic Transfer of Prokaryotic Sensors and Circuits to Mammalian Cells,” ACS Synthetic Biology, vol. 3, no. 12, pp. 880–891, 2014.

[7] W. He, S. Yuan, W. Zhong, M. A. Siddikee, and C. Dai, “Application of genetically engineered microbial whole-cell biosensors for combined chemosensing.” Applied Microbiology and Biotechnology, vol. 100, no. 3, pp. 1099–1119, 2016.

[8] L. Bousse, “Whole cell biosensors,” Sensors and Actuators B: Chemical, vol. 34, no. 1-3, pp. 270–275, 1996.

[9] Q. Gui, T. Lawson, S. Shan, L. Yan, and Y. Liu, “The Application of Whole Cell-Based Biosensors for Use in Environmental Analysis and in Medical Diagnostics,” Sensors, vol. 17, no. 7, pp. 1623, 2017.

[10] A. S. Rudolph and J. Reasor, “Cell and tissue based technologies for environmental detection and medical diagnostics,” Biosensors and Bioelectronics, vol. 16, no. 7-8, pp. 429–431, 2001.

[11] L. F. Ribeiro, T. D. Warren, and M. Ostermeier, “Construction of Protein Switches by Domain Insertion and Directed Evolution,” in Synthetic Protein Switches, vol. 1596 of Methods in Molecular Biology, pp. 43–55, Springer New York, New York, NY, 2017.

[12] V. Stein and K. Alexandrov, “Synthetic protein switches: design principles and applications,” Trends in Biotechnology, vol. 33, no. 2, pp. 101–110, 2015.

[13] N. M. Goodey and S. J. Benkovic, “Allosteric regulation and catalysis emerge via a common route,” Nature Chemical Biology, vol. 4, no. 8, pp. 474–482, 2008.

[14] D. Kern and E. R. Zuiderweg, “The role of dynamics in allosteric regulation,” Current Opinion in Structural Biology, vol. 13, no. 6, pp. 748–757, 2003.

[15] M. Shamir, Y. Bar-On, R. Phillips, and R. Milo, “SnapShot: Timescales in Cell Biology,” Cell, vol. 164, no. 6, pp. 1302–1302.e1, 2016.

[16] K. Deuschle, M. Fehr, M. Hilpert et al., “Genetically encoded sensors for metabolites,” Cytometry Part A, vol. 64A, no. 1, pp. 3–9, 2005.

[17] K. Deuschle, S. Okumoto, M. Fehr, L. L. Looger, L. Kozhukh, and W. B. Frommer, “Construction and optimization of a family of genetically encoded metabolite sensors by semirational protein engineering,” Protein Science, vol. 14, no. 9, pp. 2304–2314, 2005.

[18] M. Merkx, M. V. Golynskiy, L. H. Lindenburg, and J. L. Vinkenborg, “Rational design of RET sensor proteins based on mutually exclusive domain interactions,” Biochemical Society Transactions, vol. 41, no. 5, pp. 1201–1205, 2013.

[19] A. E. Palmer, Y. Qin, J. G. Park, and J. E. McCombs, “Design and application of genetically encoded biosensors,” Trends in Biotechnology, vol. 29, no. 3, pp. 144–152, 2011.

[20] C. M. Wright, R. C. Wright, J. R. Eshleman, and M. Ostermeier, “A protein therapeutic modality founded on molecular regulation,” Proceedings of the National Academy of Sciences of the United States of America, vol. 108, no. 39, pp. 16206–16211, 2011.

[21] B. Owens, “Smart insulin: redesign could end hypoglycemia risk,” Nature Biotechnology, vol. 36, no. 10, pp. 911–912, 2018.

[22] L. F. Ribeiro, N. Nicholes, J. Tullman et al., “Insertion of a xylanase in xylose binding protein results in a xylose-stimulated xylanase,” Biotechnology for Biofuels, vol. 8, no. 1, 2015.

[23] L. F. Ribeiro, J. Tullman, N. Nicholes et al., “A xylose-stimulated xylanase–xylose binding protein chimera created by random nonhomologous recombination,” Biotechnology for Biofuels, vol. 9, no. 1, 2016.

[24] O. Dagliyan, D. Shirvanyants, A. V. Karginov et al., “Rational design of a ligand-controlled protein conformational switch,” Proceedings of the National Academy of Sciences of the United States of America, vol. 110, no. 17, pp. 6800–6804, 2013.

[25] J. E. Dueber, “Reprogramming Control of an Allosteric Signaling Switch Through Modular Recombination,” Science, vol. 301, no. 5641, pp. 1904–1908, 2003.

[26] K. Deckert, S. J. Budiardjo, L. C. Brunner, S. Lovell, and J. Karamikolas, “Designing Allosteric Control into Enzymes by Chemical Rescue of Structure,” Journal of the American Chemical Society, vol. 134, no. 24, pp. 10055–10060, 2012.

[27] B. L. Oakes, D. C. Nadler, A. Flamholz et al., “Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch,” Nature Biotechnology, vol. 34, no. 6, pp. 646–651, 2016.

[28] L. F. Ribeiro, L. F. Ribeiro, M. Q. Barreto, and R. J. Ward, “Protein Engineering Strategies to Expand CRISPR-Cas9 Applications,” International Journal of Genomics, vol. 2018, pp. 1–12, 2018.

[29] Y. Xia, N. DiPrimio, T. R. Keppel et al., “The Designability of Protein Switches by Chemical Rescue of Structure,” Journal of the American Chemical Society, vol. 135, no. 50, pp. 18840–18849, 2013.

[30] A. Ali, J. Reis, X. Xia et al., “Optogenetic Inhibitor of the Transcription Factor CREB,” Chemistry & Biology, vol. 22, no. 11, pp. 1531–1539, 2015.

[31] R. Aroul-Selvam, T. Hubbard, and R. Sasidharan, “Domain Insertions in Protein Structures,” Journal of Molecular Biology, vol. 338, no. 4, pp. 633–641, 2004.
[96] J. M. Karchin, J. Ha, K. E. Namitz, M. S. Cosgrove, and S. N. Loh, “Small Molecule-Induced Domain Swapping as a Mechanism for Controlling Protein Function and Assembly,” *Scientific Reports*, vol. 7, no. 1, 2017.

[97] J. Lee, M. Natarajan, V. C. Nashine et al., “Surface Sites for Engineering Allosteric Control in Proteins,” *Science*, vol. 322, no. 5900, pp. 438–442, 2008.

[98] K. Reynolds, R. McLaughlin, and R. Ranganathan, “Hot Spots for Allosteric Regulation on Protein Surfaces,” *Cell*, vol. 147, no. 7, pp. 1564–1575, 2011.

[99] D. Pincus, O. Resnekov, and K. A. Reynolds, “An evolution-based strategy for engineering allosteric regulation,” *Physical Biology*, vol. 14, no. 2, p. 025002, 2017.