Chapter V: Engineering protein-responsive gene regulators and cellular biosensors

Abstract

We have previously developed small molecule-responsive RNA devices, called ribozyme switches[1], which are capable of regulating gene expression in response to target ligands and detecting the biosynthesis of target metabolites in a noninvasive manner in the cellular environment. The switch devices were constructed through a modular and portable platform comprised of distinct functional domains: the sensor, the actuator, and the transmitter. In this study, we explore the extensibility of the small molecule-responsive ribozyme-switch platform to respond to a different class of target ligands, proteins, by implementing protein-responsive RNA aptamers within the sensor domain of the device molecule. Preliminary data demonstrates that the resulting devices function as gene regulators/cellular biosensors in response to the presence of the target protein. The engineering of a protein-responsive ribozyme-switch platform may enable the construction of 'designer' gene regulators and cellular biosensors that can be used to respond to, report on, and manipulate the expression levels of a specific target protein of interest, such as those associated with a particular diseased state.

5.1. Introduction

Proteins are fundamental cellular constituents that exhibit a wide variety of functional activities and are encoded within the genomes of organisms. Numerous diseases originate from faulty and/or improperly-regulated protein expression. Genetically-encoded molecular tools are needed for controlled regulation of the expression of targeted genes and probing of corresponding protein expression. Novel gene regulators and cellular biosensors that provide noninvasive dynamic detection, control, and manipulation of target protein expression and activity are highly desired and represent powerful tools for basic and applied biological research.

Many synthetic biosensors have been developed to probe various target proteins. Antibodies are the most popular class of biosensors with a broad range of remarkable target ligand recognition capabilities[2] and have made profound contributions to the elucidation of protein interactions and functions. However, applications of antibodies as biosensors have primarily been targeted to *in vitro* studies[3, 4] or invasive studies in cellular systems[5]. In addition, the employment of antibodies in biosensing applications presents issues associated with structural stability, difficult manipulation, and use of animals for generating new antibodies[6, 7]. As a result, the development of antibody-based biosensors can be labor-intensive, time-consuming, and expensive.

Alternatively, biosensors have been developed based on fluorescent protein fusion and fluorescence resonance energy transfer (FRET) for *in vivo* studies of protein interactions and functions. In the former, proteins of interest are fused with a fluorescent protein and their expression levels and subcellular localization can be visualized using a fluorescent microscope[8]. In the latter, the synthetic genetically-encoded biosensors are composed of the

target proteins fused with appropriate donor-acceptor fluorescent protein pairs suitable for FRET[9, 10]. FRET-based biosensors have become more powerful following the development of fluorescent proteins with enhanced signal response and sensitivity[11]. Although fluorescent protein-fused biosensors are valuable tools in monitoring protein-associated cellular events in living systems, these sensors pose a couple of major challenges[9, 10]. First, many proteins are not amenable to protein fusion strategies, which may perturb the native folding state of the target protein and thus its functional activity, or the intensity of the fluorescent reporter protein. Second, there are a limited number of good FRET pairs, and many classes of protein molecules are incompatible or cannot be properly coupled with any FRET pair based on conformations and distances required for generating detectable FRET signals. Therefore, despite the tremendous value held by fluorescent protein-fused biosensors, their applicability is fairly limited to certain proteins, thus lacking platform universality and potential for broad applications. In addition, while antibody-based or fluorescent protein fusion-based biosensors are capable of probing target proteins, they are limited in application as gene regulators.

We have previously developed modular synthetic RNA devices, called ribozyme switches, based on an extensible gene-regulatory/biosensor platform that exploits RNA aptamers as sensing elements for small-molecule ligands[1]. Aptamers are nucleic acid-based molecular sensing elements that can bind ligands with high affinity and specificity[12, 13] and are well suited as sensing components of biosensors[7]. Other than antibodies, aptamers are the only molecular species that exhibit 'universal' binding activities for a diverse range of ligands[6]. The ribozyme switch platform was demonstrated to be modular such that the sensing component of the biosensor was directly replaced with a different small molecule-binding RNA aptamer, where the resulting biosensor was shown to exhibit ligand

responsiveness to this new small-molecule ligand[1]. The ability of the direct sensor replacement strategy to generate functional switches highlights the platform adaptability to construct new biosensors for various target ligands. In addition to biosensing, ribozyme switches can also function as gene regulators in response to ligand binding. However, the switch platform has not been adapted to probe proteins. Recently, a similar ribozyme-based system was demonstrated *in vitro*, in which the ribozyme cleavage activity was regulated by a protein kinase, ERK2, through molecular recognition of the kinase by its RNA aptamer coupled to the ribozyme[14]. This system, however, does not support portability to the cellular environment, as the sequence elements required for the *in vivo* functional activity of the ribozyme are absent in the design[15].

In this study, we set out to develop an *in vivo*-functional platform for the construction of RNA devices that enable both biosensing of specific protein molecules of interest and regulation of target gene expression in response to those protein molecules in living cells. We extend our small molecule-responsive ribozyme switch platform by implementing protein-binding RNA aptamers within the sensor domain of the switch molecule to generate protein-responsive ribozyme switches. Preliminary data show that the resulting switch devices function as gene regulators/cellular biosensors in response to the presence of a target protein. Additional experiments as described in Discussion will be conducted to further support and confirm the functional activity of these devices currently observed. Successful construction and demonstration of modular protein-responsive gene regulator/biosensor platform will further advance the elucidation of protein interactions and functions. In addition, such tools are essential for the construction of synthetic biological systems, which hold promise in furthering current understanding of natural biological systems.

5.2. Results

5.2.1. General composition framework and construction scheme for protein-responsive ribozyme switches

The ribozyme switch platform is comprised of three distinct functional components: a sensor component, comprised of an RNA aptamer; an actuator component, comprised of a hammerhead ribozyme including the sequence elements required for its *in vivo* cleavage activity[15]; and a transmitter component, comprised of a sequence that couples the sensor and actuator components and translates the binding event in the sensor to a conformation change in the actuator through a competitive hybridization event called strand displacement[1] (Figure 5.1). The sensor, actuator, and transmitter components are modularly coupled and functionally independent of one another.

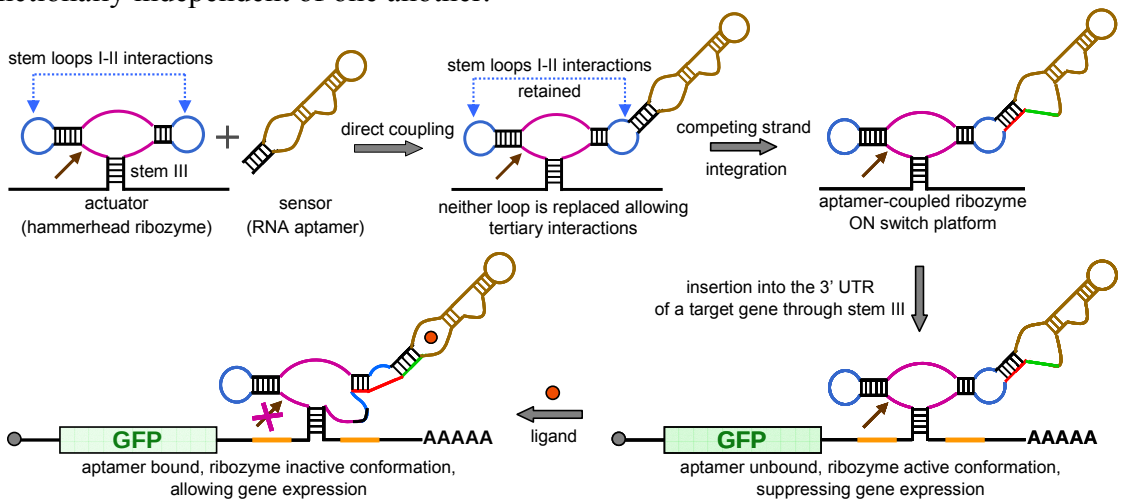

Figure 5.1. General composition framework and modular design strategy for engineering ligand-controlled ribozyme switch-based gene regulatory systems[1]. Color schemes: catalytic core, purple; loop sequences, blue; aptamer sequence, brown; competing strand, green; switching strand, red; spacer sequences, orange; cleavage site, brown arrow. An aptamer is directly attached to the ribozyme through one of its loops without replacing any part of the ribozyme, thereby maintaining loop I-II interactions required for *in vivo* functionality. Spacer sequences are included on both ends of the ribozyme switch to insulate the molecule from non-specific interactions with the surrounding sequences. A competing strand, whose sequence is similar to that of the switching strand, is integrated into the aptamer-coupled ribozyme, which enables the RNA molecule to adopt two primary conformations at equilibrium through the strand displacement mechanism. Ligand binding shifts the

equilibrium distribution towards the conformation in which the ligand-binding pocket is formed.

A general construction scheme for protein-responsive ribozyme switches is illustrated in Figure 5.2 based on a ribozyme ON switch platform, where an RNA aptamer for potentially any target protein of interest is directly replaced within the sensor domain of the switch molecule. In principle, in the absence or little expression of the target protein ligand inside the cell, the ribozyme switch will down-regulate the mRNA level within which the switch is embedded. As the expression level of the protein ligand increases and the ligand binds to its aptamer, the ribozyme switch will begin to up-regulate the level of the encoded transcript. In the case of regulating reporter gene expression, ribozyme switches can serve as cellular biosensors by transmitting a change in the levels of a protein ligand of interest to a change in the reporter gene expression level. In the case of regulating the expression of a particular target gene, ribozyme switches can function as specific protein-responsive gene regulators. The latter regulatory scheme potentially holds a therapeutic value in the treatment of protein-associated diseases.

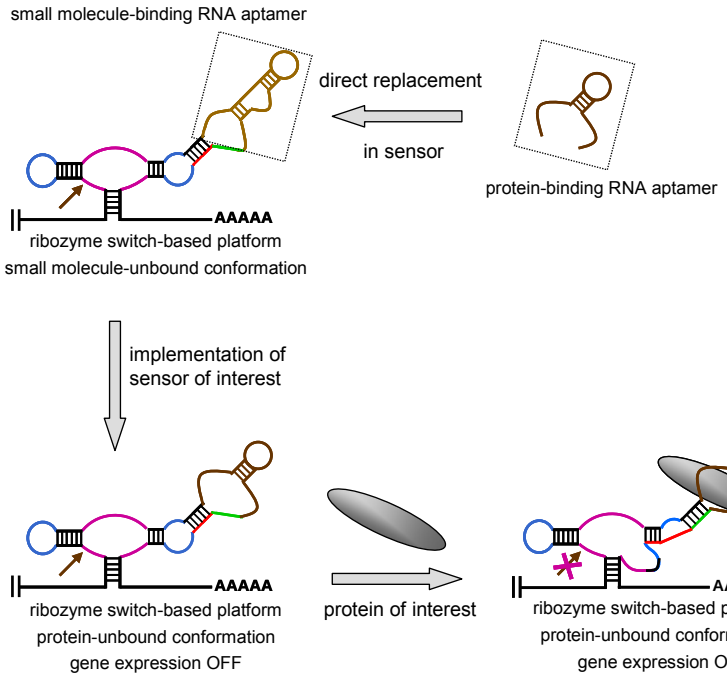

Figure 5.2. A general construction scheme for protein-responsive ribozyme switches. Color schemes follow those described in Figure 5.1.

5.2.2. Development of protein-responsive ribozyme switches

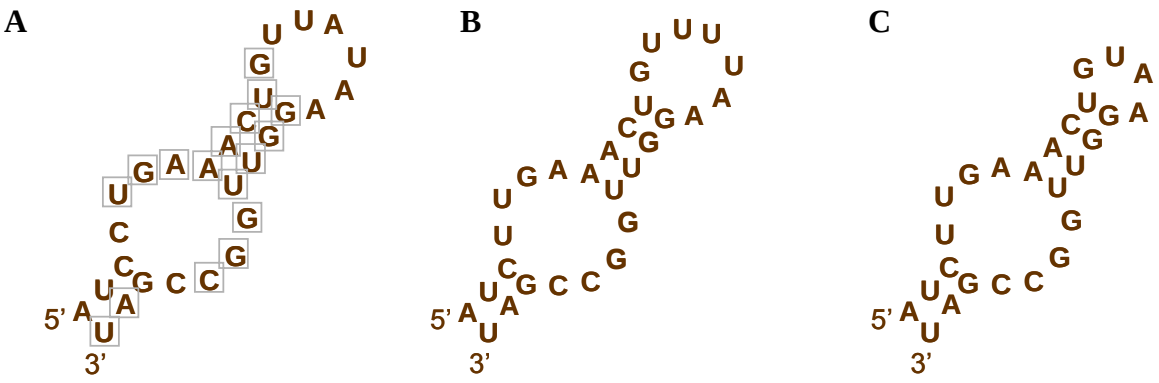

Figure 5.3. Sequences and structures of RNA aptamers with affinity and specificity for NF-κB p50. (A) the original NF-κB p50-binding RNA aptamer[16] (NF-kB1). (B) and (C) the optimized NF-κB p50-binding RNA aptamers[17, 18] (NF-kB4 and NF-kB5, respectively). The conserved nucleotides are indicated in gray boxes.

To develop a protein-responsive ribozyme switch, we employed an existing RNA aptamer with high binding affinity and specificity for the p50 subunit of the human nuclear transcription factor kappa B (NF-κB p50)[16] as a sensor component (NF-κB1, Figure 5.3A). We employed this aptamer primarily because it has been previously shown to bind its target transcription factor in a yeast three-hybrid system in activating the transcriptional event, thus supporting the *in vivo* target recognition capability of this *in vitro* selected aptamer[17, 18]. In addition to this original NF-κB p50 aptamer, we also employed two optimized derivatives of this aptamer[17, 18] (NF-κB4 and NF-κB5, respectively, Figure 5.3, B and C). We implemented these aptamers within the sensor domain of a gene ON ribozyme switch, L2bulge1[1]. As the conserved nucleotides are present in the base stem of these aptamers, we modified the transmitter domain of L2bulge1 to include these nucleotides in constructing an NF-κB p50-

responsive ribozyme switch, L2bulge1NF-κB1 (Figure 5.4). The sequences of the other two switch constructs, L2bulge1NF-κB4 and L2bulge1NF-κB5 are shown in Figure 5.5.

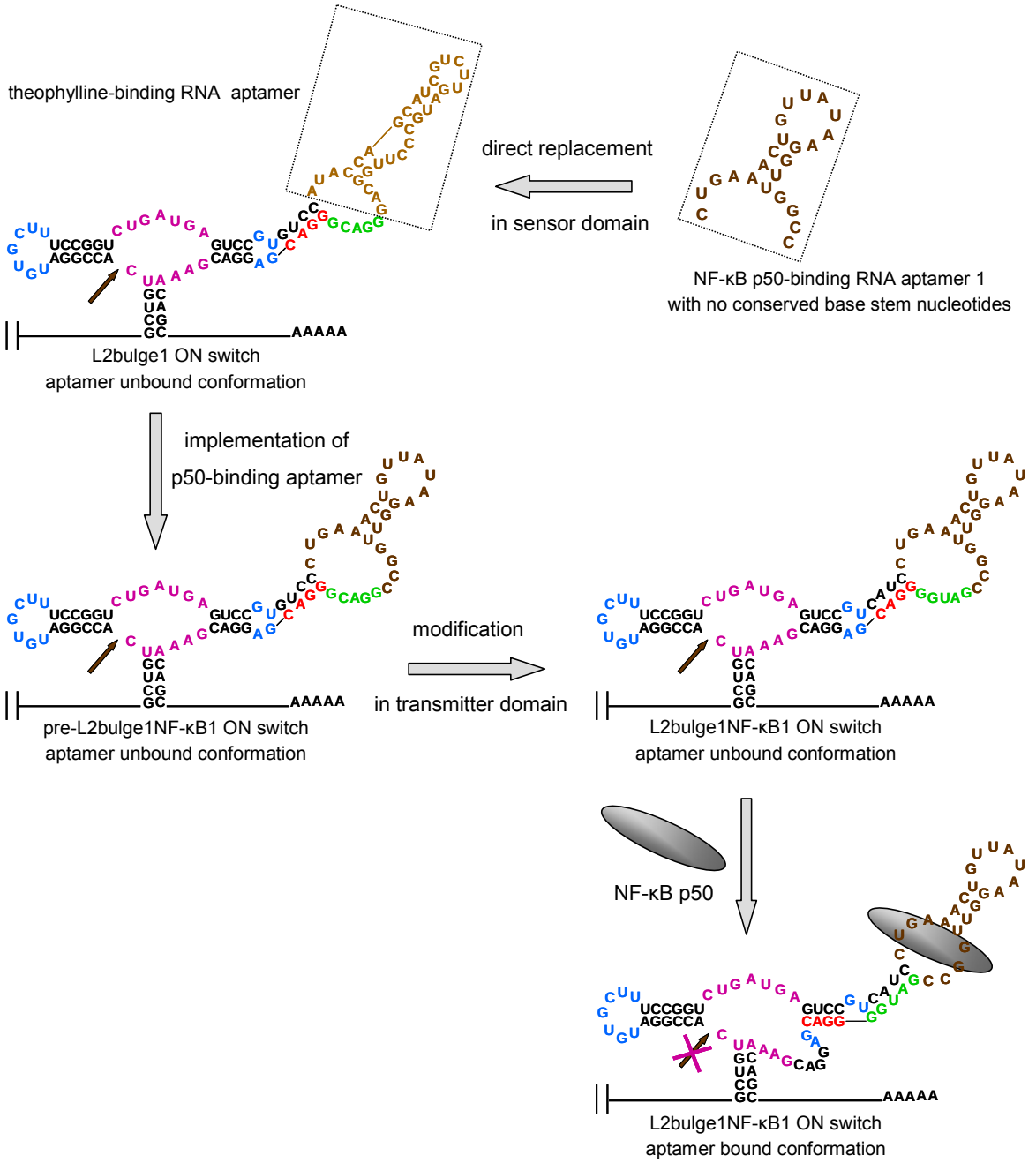

Figure 5.4. Modular design strategies and systematic engineering of an NF-κB p50-responsive ribozyme switch. Color schemes follow those described in Figure 5.1. An ON switch platform, L2bulge1, is used for illustration, where an NF-κB p50-binding aptamer (right dashed box) is directly replaced within the existing sensor domain (left dashed box) to construct a p50-responsive gene regulator/cellular biosensor. Sequence modifications were made within the transmitter domain to include the conserved nucleotides present in the base stem of the p50-binding aptamer.

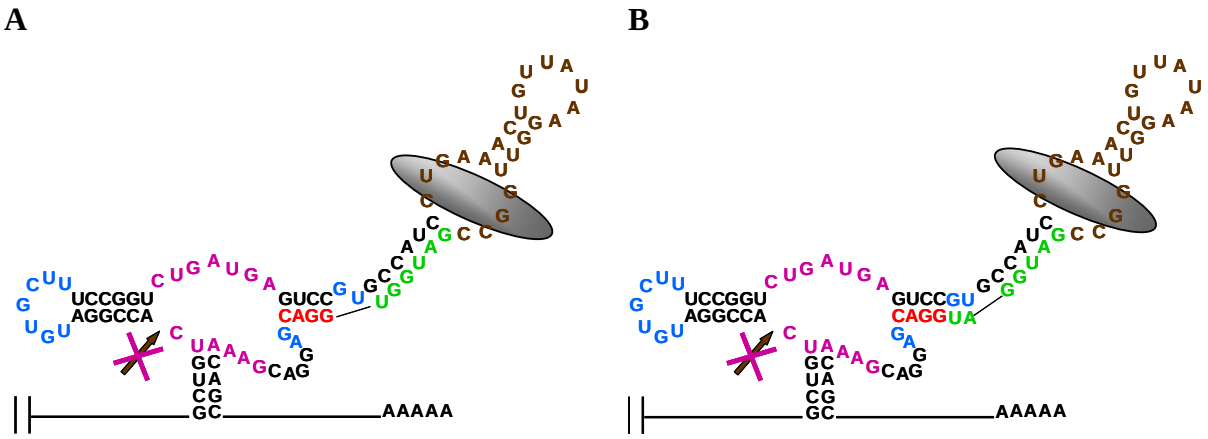

Figure 5.5. Sequences and secondary structures of two p50-responsive ribozyme switches in their ligand-bound conformations: (A) L2bulge1NF-κB4 and (B) L2bulge1NF-κB5. Color schemes follow those described in Figure 5.1. The transmitter domains of these switches are modified from that of L2bulge1 to include the conserved nucleotides present in the aptamer base stem.

5.2.3. In vivo functional activity of p50-responsive ribozyme switches

The resulting ribozyme switches were each integrated into the 3' untranslated region (UTR) of a fluorescent reporter gene and expressed from a ribozyme-switch characterization plasmid. As the plasmid employed in the yeast three hybrid system, which contains the gene encoding NF-κB p50 fused with GAL4 activation domain (GAL4AD) and tagged with a nuclear localization signal[18], was compatible with our expression system, this plasmid was co-transformed with each ribozyme switch to endogenously express the protein ligand in *Saccharomyces cerevisiae* cells. The up-regulation of reporter gene expression was observed in cells harboring both the ribozyme switch and p50 expression constructs in comparison to cells harboring just the ribozyme switch and no p50 expression construct, demonstrating the protein-responsive gene regulatory/biosensor function of these ribozyme switches (Figure 5.6).

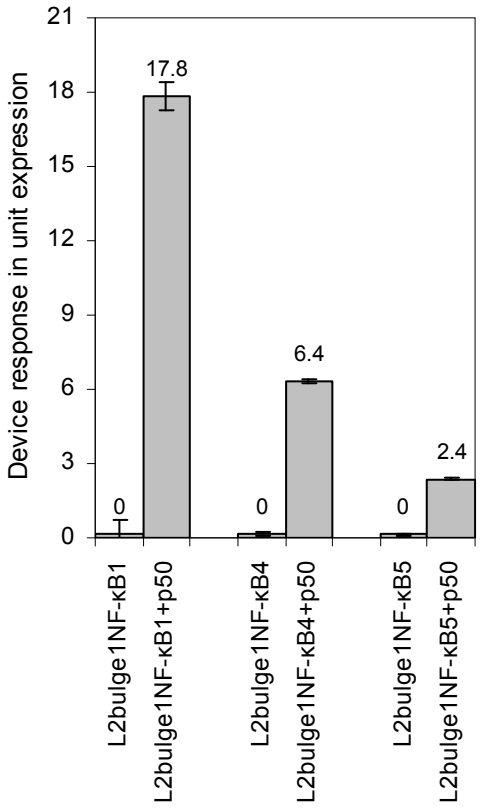

Figure 5.6. The device response of p50-responsive ribozyme switches exhibiting ON switch regulatory responses. Device response is reported in unit expression as described in Materials and Methods.

Among the three p50-responsive ribozyme switches, L2b1NF-κB1 comprised of the original *in vitro*-selected aptamer, exhibits the largest regulatory response (Figure 5.6) compared to the other two switches, L2b1NF-κB4 and L2b1NF-κB5, which are comprised of the optimized aptamer sequences, indicating that the optimization in the yeast three-hybrid system was not translatable to the ribozyme-switch system. This may be explained by the fact that these sequences were optimized in the context of the three hybrid system to more effectively activate transcription and thus might be dependent on the surrounding sequences to which they were coupled in the hybrid system to achieve the optimized activity.

5.3. Discussion and Future Work

The preliminary data show that the ribozyme switches comprised of p50-binding RNA aptamers as sensing elements are capable of regulating target gene expression through molecular recognition of p50 in living cells, thus serving as p50-responsive gene regulators and cellular biosensors. However, additional experiments will be conducted to further support and confirm the p50-responsive ribozyme-switch activity currently observed. First, the p50-directed switch activity will be characterized without GAL4AD fused into p50 to exclude the possible non-specific activity arising from the presence of GAL4AD. In addition, nuclear localization signal will be removed from p50 to examine if the switch activity is still observed when the ligand protein is present only in the cytoplasm. This study will provide insights into the mechanism of ribozyme switch regulation, specifically where the binding-cleavage events occur inside the cell. Second, p50 will be expressed under the control of an inducible promoter system so that a transfer function of device response may be generated across a gradient of p50 expression levels. Third, mutational studies will be conducted by introducing a single or few nucleotide mutations within the p50-binding aptamer sequences that are involved in p50 binding to confirm that the observed regulatory and sensing activities are indeed due to the precise recognition of the target protein. Fourth, quantification of cellular transcript levels through quantitative real-time polymerase chain reaction (qRT-PCR) will be carried out, which will provide information associated with p50-dependent regulation of transcripts through catalytic cleavage. Fifth, protein gel assay using Coomassie brilliant blue[19] or Western blots[20] will be performed to ensure *in vivo* expression of the effector protein p50. Finally, the regulation of gene expression through more advanced devices will be examined by integrating a single-input small molecule-responsive ribozyme switch and a

single-input protein-responsive ribozyme switch within the 3' UTR of a reporter gene. This implementation allows sophisticated control of gene expression in response to two different classes of input ligands, small molecules and proteins, and thus represents a higher-order signal integration scheme. The p50-binding RNA aptamers may also be implemented within the sensor domain of a gene OFF switch, such as L2bulgeOff1[1]. The functional demonstration of a protein-responsive OFF switch will broaden the application areas of these protein-responsive gene regulators/cellular biosensors.

Successful demonstration of ribozyme switch-based protein-responsive gene regulators and cellular biosensors will provide a significant advance in the current technologies available for controlling gene expression and probing protein interactions and functions. For instance, these gene regulators will enable the reprogramming of cellular behavior through regulation of the expression of certain proteins of interest in a specific effector protein-dependent manner, where the regulated protein may be different from the effector protein and the latter may be native to the cellular system employed. In regulating the expression of reporter proteins, ribozyme switches will serve as noninvasive synthetic cellular biosensors to monitor temporal and spatial fluctuations in the expression levels of both exogenous and endogenous proteins of interest. Therefore, the development of protein-responsive gene regulators and cellular biosensors based on a modular and extensible functional platform is critical for broader applications of these tools to a wide range of basic and applied biological research.

5.4. Materials and Methods

5.4.1. Plasmid construction and transformation

A ribozyme-switch characterization low-copy plasmid, pRzS-TEF1, harboring the yeast enhanced green fluorescent protein (yEGFP) under the control of a TEF1 promoter and an ADH1 terminator was used to characterize protein-responsive ribozyme switches. Ribozyme switches comprising the p50-binding RNA aptamer in the sensor domain were cloned into the 3' UTR of the *yegfp* gene, using two unique restriction sites, Avr II and Xho I, 3 nucleotides downstream of the stop codon of yEGFP. The gene encoding p50 under the control of an ADH1 promoter was expressed from a high-copy plasmid[18].

The engineered ribozyme constructs were generated by PCR amplification using the appropriate oligonucleotide templates and primers. Cloned plasmids were transformed into an electrocompetent *Escherichia coli* strain, DH10B (Invitrogen) and all ribozyme constructs were confirmed by subsequent sequencing (Laragen, Inc). Confirmed ribozyme plasmids were co-transformed with the plasmid encoding p50 into a *Saccharomyces cerevisiae* strain (W303 *MATα his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1*) using standard lithium acetate procedures[21]. As a control, each ribozyme plasmid was transformed alone into *S. cerevisiae*, which serves as the basal expression level of the ribozyme construct in the absence of p50 expression.

5.4.2. RNA secondary structure prediction and free energy calculation

RNAstructure 4.2 (http://rna.urmc.rochester.edu/rnastructure.html) was used to predict the secondary structures of all switch constructs and their thermodynamic properties as described previously[1]. RNA sequences that were predicted to adopt at least two stable equilibrium conformations (ribozyme inactive and active) were constructed and examined for functional activity.

5.4.3. Ribozyme characterization assays

Ribozyme switch activities were examined through characterization assays as described previously[1]. *S. cerevisiae* cells harboring the appropriate plasmids were grown in synthetic complete medium supplemented with an appropriate dropout solution and sugar overnight at 30^{o}C. Overnight cultures were back diluted into fresh medium to an optical density at 600 nm (OD_{600}) of approximately 0.1 and grown at 30^{o}C. Back-diluted cells were grown to an OD_{600} of 0.8-1.0 or for a period of approximately 6 h before measuring GFP levels on a Cell Lab Quanta SC flow cytometer (Beckman Coulter).

5.4.4. Fluorescence quantification

Fluorescence measurements were taken following the protocols described previously[1]. The population-averaged fluorescence of each sample was measured on a Safire fluorescence plate reader with the following settings: excitation wavelength of 485 nm, an emission wavelength of 515 nm, and a gain of 100. Fluorescence readings were normalized to cell number by dividing fluorescence units by the OD_{600} of the cell sample and subtracting the background fluorescence level to eliminate autofluorescence.

Fluorescence distributions within the cell populations were measured on a Quanta flow cytometer with the following settings: 488 nm laser line, 525 nm bandpass filter, and PMT setting of 5.83. Fluorescence data was collected under low flow rates for approximately 30,000 cells. Viable cells were selected and fluorescence levels were determined from 10,000 counts in this selected population. Background fluorescent cell population was used to set a ‘negative GFP’ gate. Cells exhibiting fluorescence above this negative gate are defined as the ‘positive GFP’ cell population.

Ligand protein-directed regulatory effects are reported as device response in units of expression normalized to the levels in the absence of ligand where 1 unit expression is defined as the reporter gene expression level of the native ribozyme, sTRSV[15], relative to the background fluoresence level. The expression level of sTRSV is ~2% of that of the inactive ribozyme or the full transcriptional range of 50 units of expression[1]. All fluorescence data and mean ±s.d. are reported from at least three independent experiments.

Acknowledgements

We thank A. Babiskin for providing pRzS and J. Liang and L. d'Espaux for experimental assistance. This work was supported by the Arnold and Mabel Beckman Foundation, the National Institutes of Health, and the Center for Biological Circuit Design at Caltech (fellowship to M.N.W.).

References

1. Win, M. N. & Smolke, C. D. From the Cover: A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. Proc Natl Acad Sci U S A 104, 14283-8 (2007).
2. Angenendt, P. Progress in protein and antibody microarray technology. Drug Discov Today 10, 503-11 (2005).
3. Backmann, N. et al. A label-free immunosensor array using single-chain antibody fragments. Proc Natl Acad Sci U S A 102, 14587-92 (2005).
4. Mattoon, D., Michaud, G., Merkel, J. & Schweitzer, B. Biomarker discovery using protein microarray technology platforms: antibody-antigen complex profiling. Expert Rev Proteomics 2, 879-89 (2005).

5. Ghaemmaghami, S. et al. Global analysis of protein expression in yeast. Nature 425, 737-41 (2003).

6. Jayasena, S. D. Aptamers: an emerging class of molecules that rival antibodies in diagnostics. Clin Chem 45, 1628-50 (1999).

7. O'Sullivan, C. K. Aptasensors--the future of biosensing? Anal Bioanal Chem 372, 44-8 (2002).

8. Huh, W. K. et al. Global analysis of protein localization in budding yeast. Nature 425, 686-91 (2003).

9. Hahn, K. & Toutchkine, A. Live-cell fluorescent biosensors for activated signaling proteins. Curr Opin Cell Biol 14, 167-72 (2002).

10. Li, I. T., Pham, E. & Truong, K. Protein biosensors based on the principle of fluorescence resonance energy transfer for monitoring cellular dynamics. Biotechnol Lett 28, 1971-82 (2006).

11. Nguyen, A. W. & Daugherty, P. S. Evolutionary optimization of fluorescent proteins for intracellular FRET. Nat Biotechnol 23, 355-60 (2005).

12. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249, 505-10 (1990).

13. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. Nature 346, 818-22 (1990).

14. Vaish, N. K. et al. Monitoring post-translational modification of proteins with allosteric ribozymes. Nat Biotechnol 20, 810-5 (2002).

15. Khvorova, A., Lescoute, A., Westhof, E. & Jayasena, S. D. Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. Nat Struct Biol 10, 708-12 (2003).

16. Lebruska, L. L. & Maher, L. J., 3rd. Selection and characterization of an RNA decoy for transcription factor NF-kappa B. Biochemistry 38, 3168-74 (1999).

17. Cassiday, L. A. & Maher, L. J., 3rd. In vivo recognition of an RNA aptamer by its transcription factor target. Biochemistry 40, 2433-8 (2001).

18. Cassiday, L. A. & Maher, L. J., 3rd. Yeast genetic selections to optimize RNA decoys for transcription factor NF-kappa B. Proc Natl Acad Sci U S A 100, 3930-5 (2003).

19. Uefuji, H., Ogita, S., Yamaguchi, Y., Koizumi, N. & Sano, H. Molecular cloning and functional characterization of three distinct N-methyltransferases involved in the caffeine biosynthetic pathway in coffee plants. Plant Physiol 132, 372-80 (2003).

20. Mizuno, K. et al. Isolation of a new dual-functional caffeine synthase gene encoding an enzyme for the conversion of 7-methylxanthine to caffeine from coffee (Coffea arabica L.). FEBS Lett 534, 75-81 (2003).

21. Gietz, R. & Woods, R. in Guide to Yeast Genetics and Molecular and Cell Biology, Part B, (eds. Guthrie, C. & Fink, G.) 87-96 (Academic Press, San Diego, 2002).