Chapter III: A modular and extensible RNA-based gene-regulatory platform for engineering cellular function*

Abstract

Engineered biological systems hold promise in addressing pressing human needs in chemical processing, energy production, materials construction, and maintenance and enhancement of human health and the environment. However, significant advancements in our ability to engineer biological systems have been limited by the foundational tools available for reporting on, responding to, and controlling intracellular components in living systems. Portable and scalable platforms are needed for the reliable construction of such communication and control systems across diverse organisms. We report an extensible RNA-based framework for engineering ligand-controlled gene regulatory systems, called ribozyme switches, that exhibit tunable regulation, design modularity, and target specificity. These switch platforms contain a sensor domain, comprised of an aptamer sequence, and an actuator domain, comprised of a hammerhead ribozyme sequence. We examined two modes of standardized information transmission between these domains and demonstrate a mechanism that allows for the reliable and modular assembly of functioning synthetic RNA switches and regulation of ribozyme activity in response to various effectors. In addition to demonstrating the first examples of small molecule-responsive, *in vivo* functional allosteric hammerhead ribozymes, this work describes a general approach for the construction of portable and scalable gene-regulatory systems. We demonstrate the versatility of the platform in implementing application-specific control systems for small molecule-mediated regulation of cell growth and non-invasive *in vivo* sensing of metabolite production.

*Reproduced with permission from: M. N. Win and C. D. Smolke. (2007) *Proc. Natl. Acad. Sci. U. S. A.*, 104, 14283-14288.

3.1. Introduction

Basic and applied biological research and biotechnology are limited by our ability to get information into and out from living systems, and to act on information inside living systems[1, 2]. For example, there are only a small number of inducible promoter systems available to provide control over gene expression in response to exogenous molecules[3, 4]. Many of the molecular inputs to these systems are not ideal for broad implementation, as they can be expensive and introduce undesired pleiotropic effects. In addition, broadly-applicable methods for getting information out of cells non-invasively has been limited to strategies that rely on protein and promoter fusions to fluorescent proteins, which enable researchers to monitor protein levels and localization and transcriptional outputs of networks, leaving a significant amount of the cellular information content currently inaccessible.

To address these challenges scalable platforms are needed for reporting on, responding to, and controlling any intracellular component in a living system. A striking example of a biological communication and control system is the class of RNA regulatory elements called riboswitches, comprised of distinct sensor and actuation (gene regulatory) functions, that control gene expression in response to specific ligand concentrations[5]. Building on these natural examples, engineered riboswitch elements have been developed for use as synthetic ligand-controlled gene regulatory systems[6-9]. However, these early examples of riboswitch engineering do not address the challenges posed above because they lack portability across organisms and systems, and their designs and construction do not support modularity and component reuse.

We set out to develop a universal and extensible RNA-based platform that will provide a framework for the reliable design and construction of gene regulatory systems that

can control the expression of specific target genes in response to various effector molecules. We implemented five engineering design principles (DPs) in addressing this challenge: *scalability* (DP1: a sensing platform enabling *de novo* generation of ligand-binding elements for implementation within the sensor domain); *portability* (DP2: a regulatory element that is independent of cell-specific machinery or regulatory mechanisms for implementation within the actuator domain); *utility* (DP3: a mechanism through which to modularly couple the control system to functional level components); *composability* (DP4: a mechanism by which to modularly couple the actuator and sensor domains without disrupting the activities of these individual elements); and *reliability* (DP5: a mechanism through which to standardize the transmission of information from the sensor domain to the actuator domain). A glossary of terms is available in Supplementary Text 3.1.

3.2. Results

3.2.1. Component specification for a scalable and portable gene-regulatory system

To satisfy the engineering design principle of scalability (DP1) we chose RNA aptamers[10], nucleic acid ligand-binding molecules, as the sensing platform for the universal control system. Our choice of sensing platform was driven by the proven versatility of RNA aptamers. Standard *in vitro* selection strategies or SELEX[11, 12] have been used to generate RNA aptamers *de novo* to a wide variety of ligands, including small molecules, peptides, and proteins[13]. In addition, the specificity and affinity of an aptamer can be tuned through the selection process to meet the specific performance requirements of a given application. The continued selection of new aptamers to appropriate cellular molecules that function under *in*

vivo conditions will enable these elements to be implemented as sensors in RNA-based control systems.

To satisfy the engineering design principle of portability (DP2) we chose the hammerhead ribozyme, a catalytic RNA, as the regulatory element in the universal control system. Our choice of regulatory element was driven by the ability of the hammerhead ribozyme to exhibit self-cleavage activity across various organisms and its demonstrated potential in biomedical and biotechnological applications owing to its small size, relative ease of design, and rapid kinetics[14]. The utility of hammerhead ribozymes as gene regulatory elements has been demonstrated in various systems[15-17]. In addition, several research groups have engineered a special class of synthetic hammerhead ribozymes referred to as allosteric hammerhead ribozymes that contain separate catalytic and ligand-binding domains, which interact in a ligand-dependent manner to control the activity of the ribozyme[18-21]. While this class of ribozymes enables a better control system due to the presence of the integrated ligand-binding domain, there has been no success in translating them to *in vivo* environments.

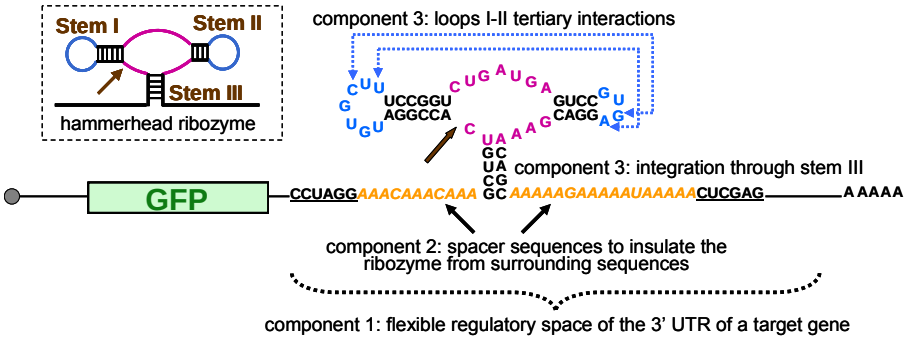
3.2.2. Design strategies for engineering portability, utility, and composability into a biological control system

To support a framework for engineering ligand-controlled gene regulatory systems, we specified a design strategy that is in accordance with our engineering principles stated above (Figure 3.1, A and B). This strategy is comprised of three components that address mechanisms for the portability (DP2), utility (DP3), and composability (DP4) of the control system and are critical to the development of a general ribozyme switch platform. First, the

cis-acting hammerhead ribozyme constructs are integrated into the flexible and portable regulatory space of the 3' UTR (Figure 3.1A). We chose to locate the synthetic ribozymes within the 3' UTR of their target gene as opposed to the 5' UTR in order to isolate their specific cleavage effects on transcript levels from their non-specific structural effects on translation initiation, as secondary structures have been demonstrated to repress efficient translation when placed in the 5' UTR[22]; K. Hawkins and C.D.S., unpublished observations). In addition, cleavage within the 3' UTR is a universal mechanism for transcript destabilization in eukaryotic and prokaryotic organisms. Second, each ribozyme construct is insulated from surrounding sequences, which may disrupt its structure and therefore its activity, by incorporating spacer sequences immediately 5' and 3' of stem III (Figure 3.1A). By implementing these two components, we ensure that these control systems will be portable across organisms and modular to coupling with different coding regions (Y. Chen and C.D.S., manuscript in preparation). The third component was necessitated by the fact that previous engineered *in vitro* allosteric ribozyme systems, which replace stem loops I or II with part of the aptamer domain (Figure 3.1B, lower right), do not function *in vivo*. From previous studies on the satellite RNA of tobacco ringspot virus (sTRSV) hammerhead ribozyme[16], we suspect that this lack of *in vivo* functionality in earlier designs results from removal of stem loop sequences that may play a critical role in tertiary interactions that stabilize the catalytically active conformation under physiological Mg^{2+} concentrations. To develop ribozyme switches that function *in vivo*, we chose to integrate the hammerhead ribozyme into the target transcript through stem III and couple the sensor domain directly to the ribozyme through stem loops I or II to maintain these potentially essential sequence elements (Figure 3.1B, upper right). We constructed a series of ribozyme controls

(Supplementary Text 3.2 and Supplementary Figure 3.1), which consist of loop coupling and stem integration controls. Implementation and characterization of the gene regulatory activity of these ribozyme constructs within a modular plasmid system in the eukaryotic model organism *Saccharomyces cerevisiae* (Figure 3.1A) indicate that maintenance of loop I and II sequences and thus integration through stem III are essential for *in vivo* functionality of such gene regulatory elements (Supplementary Figure 3.1D).

A



B

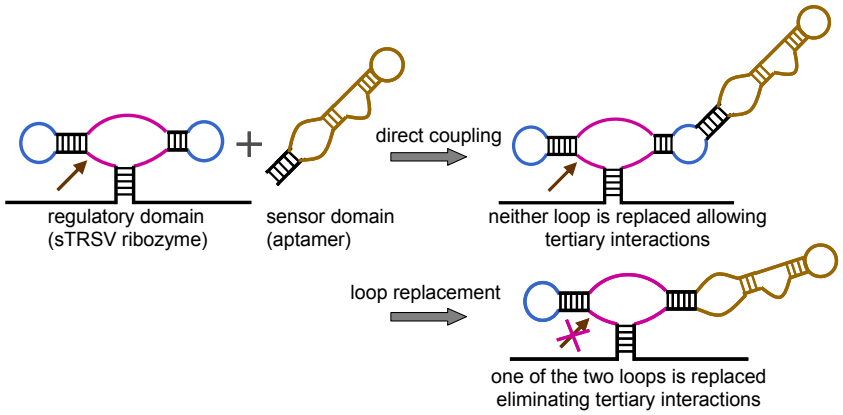


Figure 3.1. General design strategy for engineering ribozyme switches. Color schemes: catalytic core, purple; aptamer sequences, brown; loop sequences, blue; spacer sequences, yellow; brown arrow, cleavage site. (A) General compositional framework and design strategy for engineering universal cis-acting hammerhead ribozyme-based regulatory systems; restriction enzyme sites are underlined. (B) Modular coupling strategies of the sensor and regulatory domains to maintain *in vivo* activity of the individual domains.

3.2.3. Engineering mechanisms for information transmission between the modular switch domains

The final design challenge in building a universal switch platform is to develop a standardized means of transmitting information (encoded within an information transmission domain) from the sensor (aptamer) domain to the regulatory (ribozyme) domain (DP5). There are two different strategies for transmitting information between the aptamer and ribozyme domains: strand displacement and helix slipping. We constructed and characterized ribozyme switch platforms based on both mechanisms.

The first information transmission domain that we developed is based on a strand displacement mechanism, which involves the rational design of two sequences that compete for binding to a general transmission region (the base stem of the aptamer) (Figure 3.2, A and B). We employed this mechanism in engineering a ribozyme switch platform that enables both up- and down-regulation of gene expression in response to increasing effector concentrations ('ON' and 'OFF' switches, respectively). An initial ribozyme switch, L2bulge1, was constructed to up-regulate gene expression through the corresponding base platform (L2Theo, Supplementary Figure 3.1C) by incorporating a competing strand following the 3' end of the theophylline aptamer[23] (Figure 3.2A). This competing strand is perfectly complementary to the base stem of the aptamer at the 5' end. Using the same design principles, we engineered another ribozyme switch, L2bulgeOff1 (Figure 3.2B), for down-regulating gene expression. Our strand displacement strategy is based on the conformational dynamics characteristic of RNA molecules that enables them to distribute between at least two different conformations at equilibrium: one conformation in which the competing strand is not base-paired or base-paired such that the ligand-binding pocket is not formed, and the

other conformation in which the competing strand is base-paired with the aptamer base stem, displacing the switching strand and thus allowing the formation of the ligand-binding pocket. Strand displacement results in the disruption (L2bulge1) or restoration (L2bulgeOff1) of the ribozyme's catalytic core. Binding of theophylline to the latter conformation shifts the equilibrium distribution between these two conformations to favor the aptamer-bound form as a function of increasing theophylline concentration.

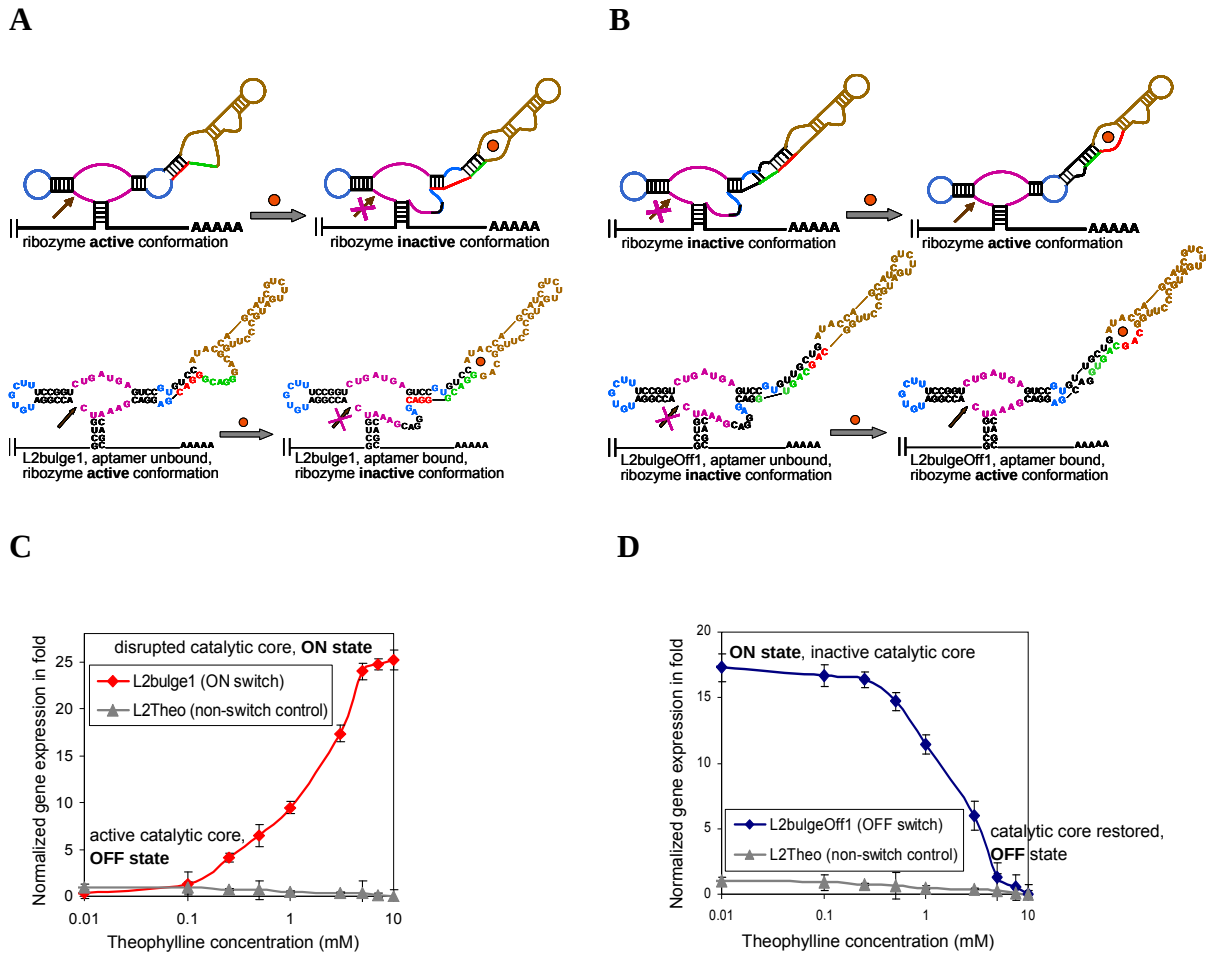


Figure 3.2. Regulatory properties of the strand displacement information transmission mechanism. Color schemes: switching strand, red; competing strand, green; all other schemes correspond to those used in Figure 3.1. (A) Gene expression 'ON' ribozyme switch platform, L2bulge1. (B) Gene expression 'OFF' ribozyme switch platform, L2bulgeOff1. The theophylline-dependent gene regulatory behavior of (C) L2bulge1 ('ON' switch), (D) L2bulgeOff1 ('OFF' switch), and L2Theo (non-switch control). Gene expression levels are

reported in fold as defined in Materials and Methods and normalized to the expression levels in the absence of effector.

Increased binding of theophylline to L2bulge1 resulted in an approximate 25 folds increase in target expression levels at 5 mM theophylline relative to those in the absence of effector (Figure. 3.2C and Supplementary Figure 3.2). In contrast, increased binding of theophylline to L2bulgeOff1 resulted in an approximate 18 folds reduction in expression levels at 5 mM theophylline relative to those in the absence of effector (Figure. 3.2D and Supplementary Figure 3.2). Through our strand displacement mechanism, we have engineered ribozyme switches *de novo*, L2bulge1 and L2bulgeOff1, that provide allosteric regulation of gene expression and function as 'ON' and 'OFF' switches, respectively.

We engineered a second class of ribozyme switch platforms to examine an alternative information transmission domain based on a helix slipping mechanism, which does not allow for rational design (Figure. 3.3A). This mechanism involves the functional screening of 'communication modules'[19-21] within the base stem of the aptamer. Communication modules are dynamic elements capable of transmitting the binding state of an aptamer domain to an adjacent regulatory domain through a 'slip-structure' mechanism[19], in which a nucleotide shift event within the element is translated to a small-scale change in the conformation of the regulatory domain in a ligand-dependent manner. These elements have been developed through *in vitro* screening processes, and their dynamic and communicative properties have been demonstrated *in vitro* in engineered allosteric ribozymes[18-21]. We screened the *in vivo* functionality of previously *in vitro* selected communication modules[19-21] by assaying the activity of these sequences within L1Theo and L2Theo. A critical difference between the design of the previously developed *in vitro* allosteric ribozymes, from which these

communication modules were generated, and that of our engineered ribozyme switches is the coupling strategies between the aptamer and ribozyme domains and their effects on the *in vivo* activity of the ribozyme domain as described previously (Figure 3.1B).

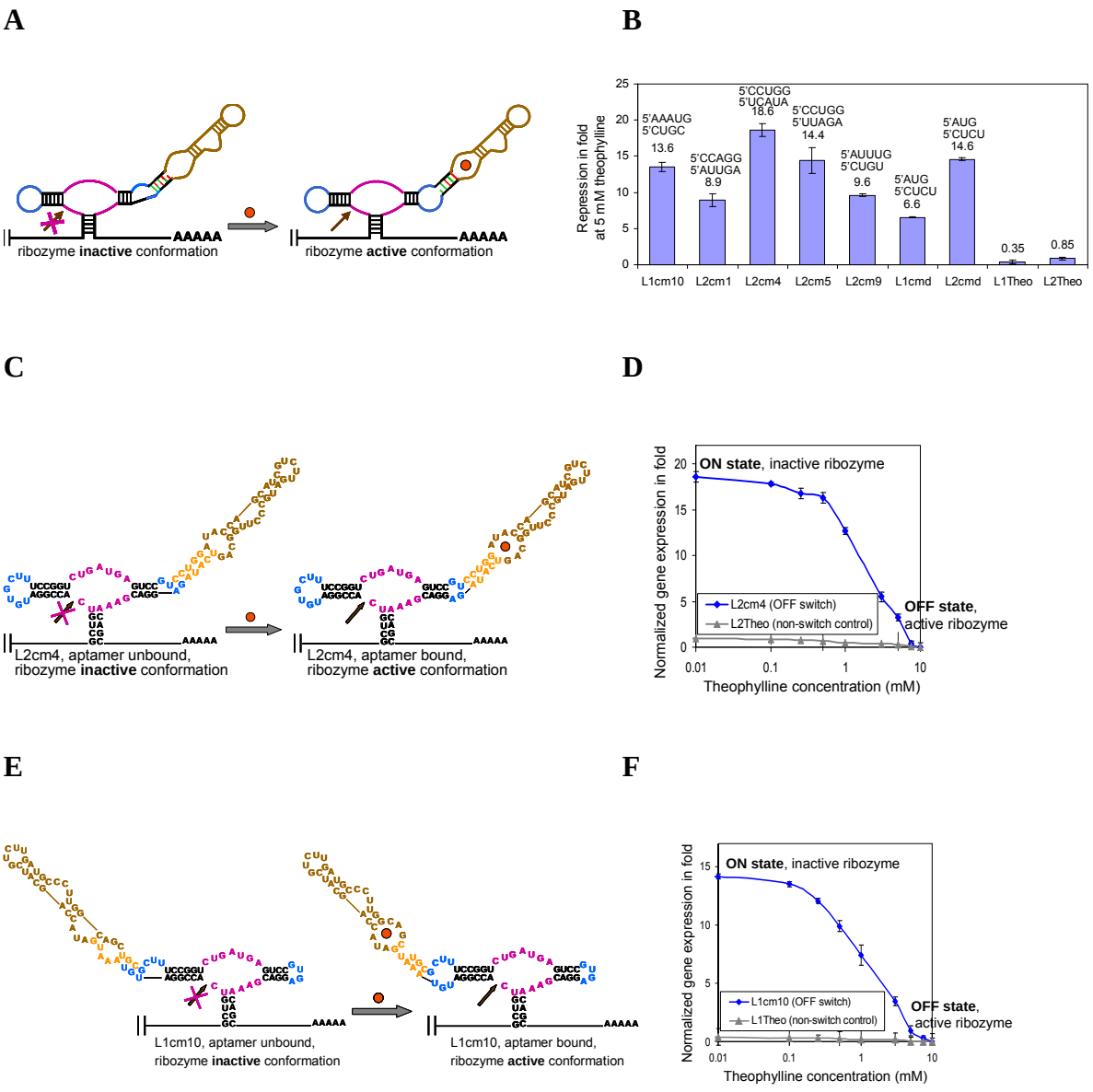


Figure 3.3. Regulatory properties of the helix slipping information transmission mechanism. Color schemes: communication module schematic, red and green; communication module sequence, orange; all other schemes correspond to those used in Figure 3.1. (A) Gene expression 'OFF' ribozyme switch platform based on helix slipping. The base stem of the aptamer is replaced with a communication module. (B) Regulatory activities of helix slipping-based ribozyme switches. Gene regulatory effects of the 'OFF' switches at 5 mM

theophylline are reported in fold repression relative to expression levels in the absence of effector. The corresponding communication module sequences are indicated. Sequence and structure of representative helix slipping ribozyme switches, (C) L2cm4 and (E) L1cm10. The theophylline-dependent gene regulatory behavior of (D) L2cm4 and (F) L1cm10. Gene expression levels are reported as described in Figure 3.2, except that in (F) L1Theo is used as a non-switch control.

Among the thirteen communication modules[19-21] screened for *in vivo* activity, five (cm1, cm4, cm5, cm9, and cmd) exhibit down-regulation of expression levels through loop II, whereas only two (cm10 and cmd) exhibit such regulation through loop I (Figure 3.3B). The regulatory activities of two helix slipping-based ribozyme switches, L2cm4 (Figure 3.3, C and D, and Supplementary Figure 3.3) and L1cm10 (Figure 3.3, E and F, and Supplementary Figure 3.3), were characterized across a range of theophylline concentrations and exhibit substantial regulatory effects. Although the helix slipping constructs are comprised of identical aptamer and catalytic core sequences, they exhibit different extents of regulation. This variability suggests that each construct contains a different equilibrium distribution between the adoptable conformations and that the energy required for structural switching between the conformations is also different.

We validated the regulatory mechanisms of representative strand displacement- and helix slipping-based switches. Relative steady-state transcript levels in the absence and presence of effector are consistent with corresponding fluorescent protein levels (Supplementary Table 3.1), indicating that cleavage in the 3' UTR results in rapid decay and inactivation of the target transcript. In addition, we demonstrated that changes in expression levels are induced shortly after effector addition (Supplementary Figure 3.4), indicating that the response of the regulatory elements to changes in effector levels is relatively rapid.

3.2.4. Rational tuning strategies enable programming of switch regulatory response

The ability to program the regulatory response of a universal switch platform is an important property in tuning the platform performance to comply with the design specifications for a particular application. We demonstrate that our strand displacement-based ribozyme switch platform incorporates an information transmission mechanism that is amenable to rational tuning strategies for programming regulatory response properties. Programming of new regulatory information is achieved by sequence alteration resulting in a change in the molecule's structural stability, which may affect its conformational switching dynamics if the molecule can adopt multiple conformations. These rational sequence modification tuning strategies are not applicable to communication module-based switches due to an inability to predict their activities.

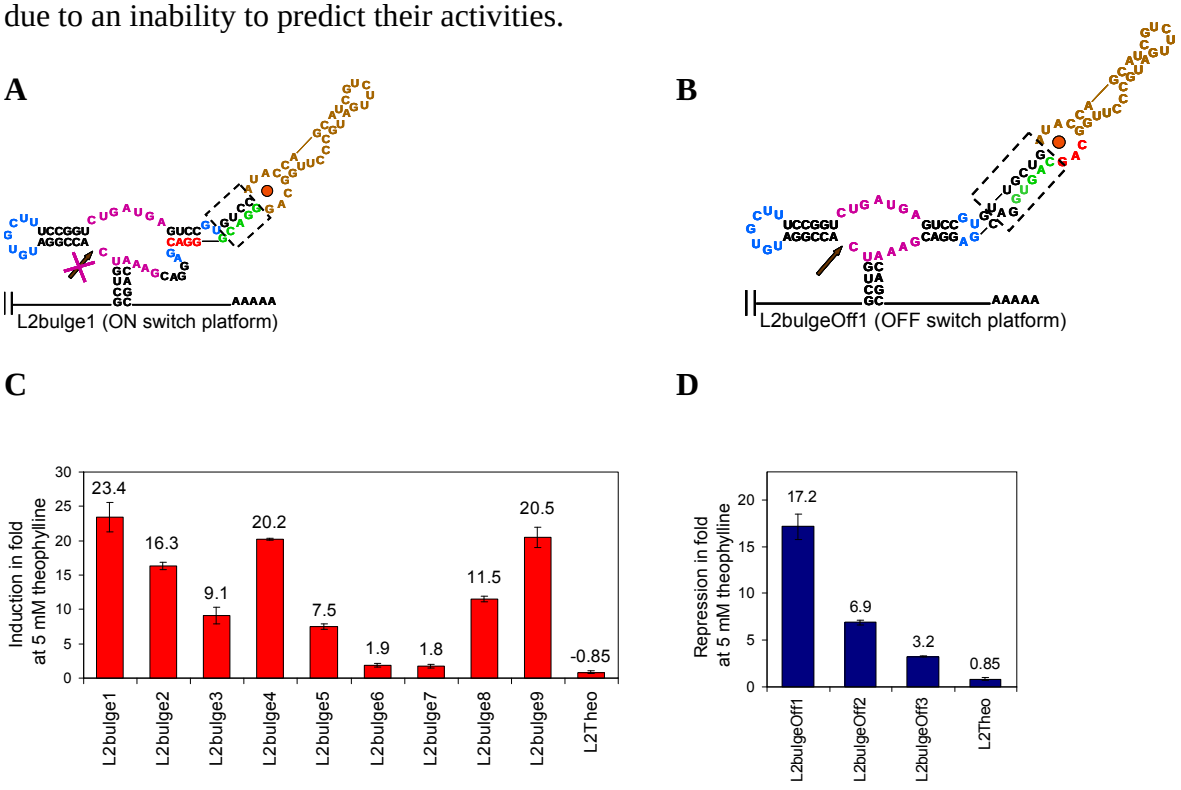


Figure 3.4. Tunability of the strand displacement-based ribozyme switches. Sequences targeted by the rational tuning strategies are indicated in the dashed boxes on the effector-bound conformations of (A) L2bulge1 (ribozyme inactive) and (B) L2bulgeOff1 (ribozyme

active). Regulatory activities of tuned strand displacement-based (C) 'ON' and (D) 'OFF' ribozyme switches. Gene regulatory effects of these switches at 5 mM theophylline are reported in fold induction for 'ON' switches and fold repression for 'OFF' switches relative to the expression levels in the absence of theophylline as described in Figure 3.2.

A more complete description of our tuning strategies is provided in Supplementary Text 3.3, Supplementary Figure 3.5, and Supplementary Table 3.2. Briefly, our rational tuning strategies target alteration of the nucleotide composition of the base stem of the aptamer domain to affect the stabilities of individual switch constructs and the energies required for the construct to switch between two adoptable conformations. Using these strategies, we rationally engineered a series of tuned 'ON' and 'OFF' switches from L2bulge1 and L2bulgeOff1, respectively (Figure 3.4, A and B). These tuned switches exhibit different regulatory ranges in accordance with our rational energetic tuning strategies (Figure 3.4, C and D, and Supplementary Figure 3.6).

3.2.5. The ribozyme switch platform exhibits component modularity and specificity

In implementing a standardized mechanism through which to transmit information between the domains of a switch platform (DP5), we needed to confirm that the modular coupling between the aptamer and ribozyme domains is maintained (DP4). We performed modularity studies on our strand displacement-based ribozyme switch platform, in which aptamers possessing sequence flexibility in their base stems can be swapped into the sensor domain. To begin to demonstrate that ribozyme switch activity may be controlled by different effector molecules we replaced the theophylline aptamer of L2bulge1 with a tetracycline mini-aptamer[24] to construct a tetracycline-repsonsive ON switch (L2bulge1tc) (Figure 3.5A). Despite similar aptamer ligand affinities[23, 24], the extent of up-regulation with

L2bulge1tc was greater than that with L2bulge1 at the same extracellular concentration of their respective ligands (Figure 3.5B). This is likely due to the high cell permeability of tetracycline[25] compared to theophylline[26]. These results demonstrate that our strand displacement-based ribozyme switch platform maintains modularity between the aptamer and ribozyme domains. We also performed similar modularity studies on the helix slipping-based switch platform by replacing the theophylline aptamer of L1cm10, L2cm4 and L2cm5 with the tetracycline mini-aptamer (L1cm10, L2cm4tc, and L2cm5tc, respectively). These constructs do not exhibit effector-mediated gene-regulatory effects (data not shown).

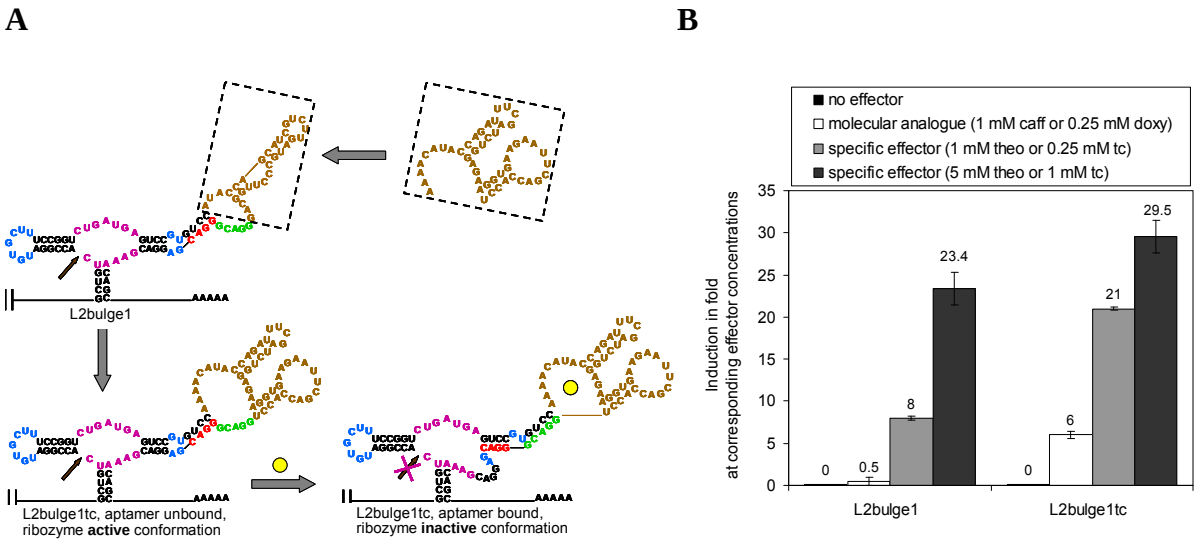


Figure 3.5. Modularity and specificity of the strand displacement-based ribozyme switches. (A) Modular design strategies for the construction of new ribozyme switches. The theophylline (left dashed box) and tetracycline (right dashed box) aptamers are shown. (B) Regulatory activities of the modular ribozyme switch pair, L2bulge1 and L2bulge1tc, in response to their respective ligands, theophylline (theo) and tetracycline (tc), and closely-related analogues, caffeine (caff) and doxycycline (doxy). Regulatory effects are reported in fold induction relative to the expression levels in the absence of effector as described in Figure 3.2.

We also demonstrated that the aptamer sequences (theophylline and tetracycline) incorporated into our ribozyme switch platforms maintain highly specific target recognition capabilities *in vivo* similar to their *in vitro* specificities generated during the selection process

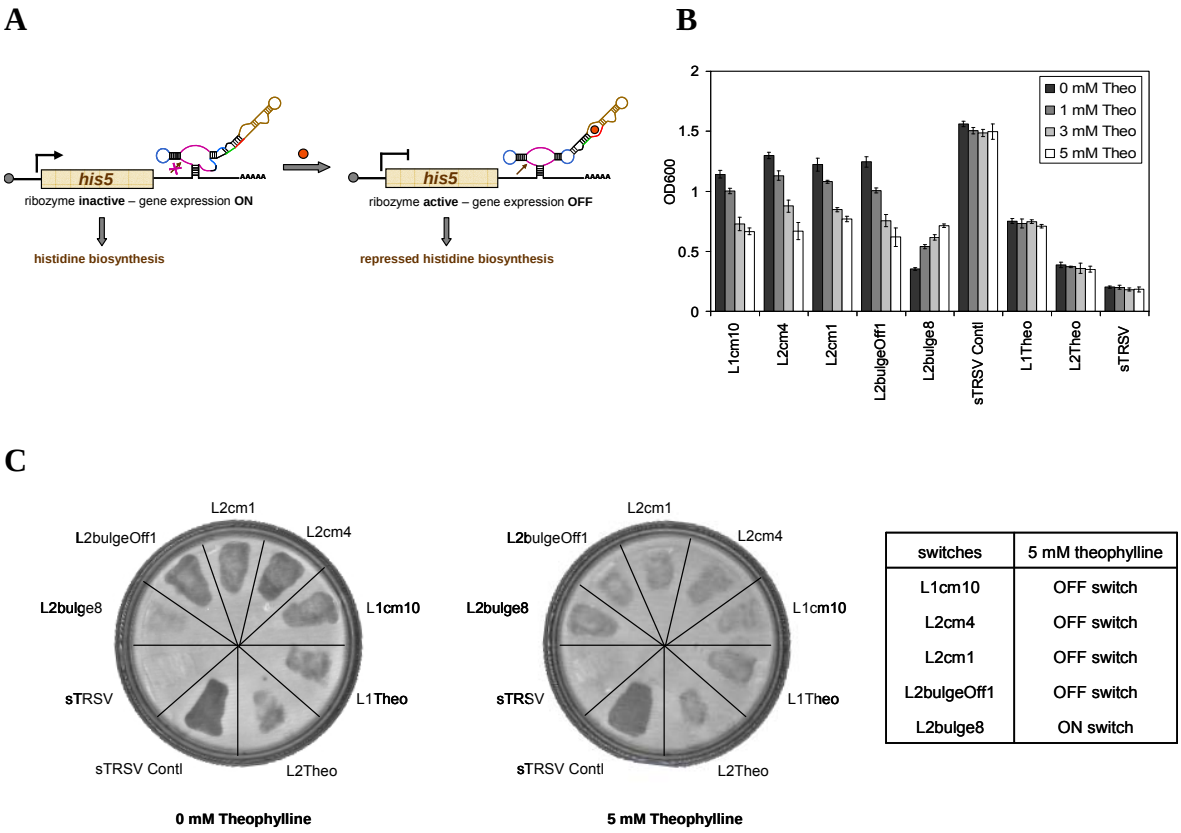
against corresponding molecular analogues (caffeine and doxycycline, respectively)[23, 24] (Figure 3.5B). This is an important property in implementing these platforms in cellular engineering applications that involve complex environments where molecular species similar to the target ligand may be present.

3.2.6. Component modularity enables implementation of ribozyme switches as regulatory systems in diverse applications

To demonstrate the scalability and utility of these switch platforms as application-specific control systems, we demonstrate the implementation of ribozyme switches in two distinct cellular engineering application areas. First, utility (DP3) and the ability to respond to and control cellular information is demonstrated by the application of ribozyme switches to small molecule-mediated regulation of cell growth. Second, scalability (DP1) and the ability to respond to and report on cellular information is demonstrated by the implementation of ribozyme switches as non-invasive *in vivo* sensors of metabolite production.

The first system explores the application of our ribozyme switches to the regulation of a survival gene, where modification of expression levels is expected to produce an observable and titratable phenotypic effect on cell growth. The reporter gene within the original constructs was replaced with a growth-associated gene (*his5*) responsible for the biosynthesis of histidine in yeast[27] (Figure 3.6A). We performed growth regulation assays across various effector concentrations using representative switch constructs and demonstrated that these switches mediate cell growth in a highly effector-dependent manner (Figure 3.6B). Plate-based assays confirm the theophylline-dependent ribozyme switch-based

regulation of cell growth (Figure 3.6C). This application demonstrates the utility (DP3) of our switch platform, in which the control system exhibits modularity to the functional level components in the regulatory system.



switches	5 mM theophylline
L1cm10	OFF switch
L2cm4	OFF switch
L2cm1	OFF switch
L2bulgeOff1	OFF switch
L2bulge8	ON switch

Figure 3.6. System modularity of ribozyme switches enables implementation in programmed cell growth. (A) System design for ribozyme switch-based regulation of cell growth. Small molecule-mediated regulation of a gene required for cell growth is illustrated for a strand displacement-based 'OFF' switch. (B) Theophylline-mediated ribozyme switch-based regulation of cell growth. Changes in growth are reported as OD_{600} values for cells grown in 5 mM 3AT in media lacking histidine. (C) Demonstration of theophylline-regulated cell growth by ribozyme switches through plate-based assays. Cells harboring ribozyme switches and control constructs were streaked on two plates containing the same medium except different effector concentrations (0 mM versus 5 mM theophylline). OFF switches (L1cm10, L2cm4, L2cm1, L2bulgeOff1) exhibit suppressed cell growth on the plate containing 5 mM theophylline while an ON switch (L2bulge8) exhibits a higher growth level on the plate containing 5 mM theophylline. The control constructs (L1Theo, L2Theo, sTRSV Contl, and sTRSV) exhibit similar growth levels on both plates. sTRSV exhibits no cell growth due to its efficient cleavage activity and sTRSV Contl exhibits the highest levels of growth due to its lack of cleavage activity.

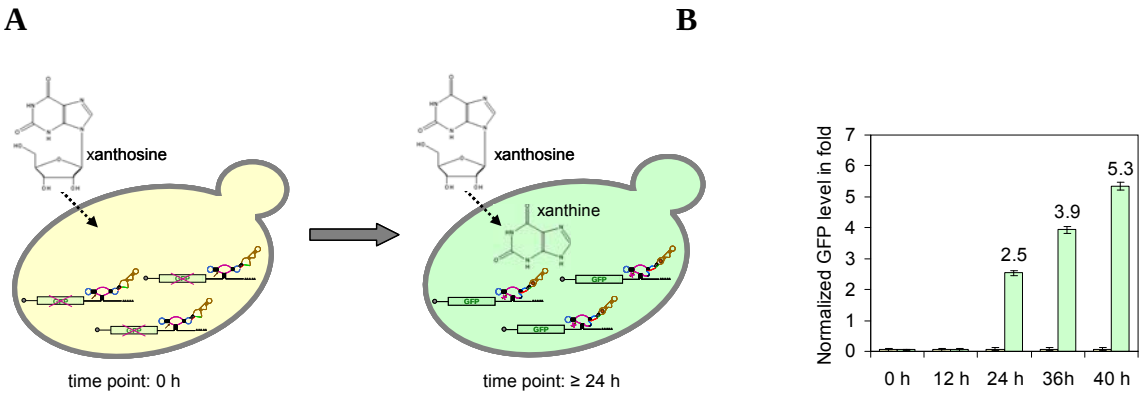


Figure 3.7. System modularity of ribozyme switches enables implementation in non-invasive detection of metabolite biosynthesis. (A) System design for ribozyme switch-based *in vivo* sensing of metabolite production. Xanthine is converted from fed xanthosine through an activity endogenous to yeast and product accumulation over time is detected through a strand displacement-based xanthine-responsive 'ON' switch coupled to the regulation of a reporter protein. (B) Ribozyme switch-based xanthine synthesis detection through L2bulge9. Metabolite sensing through L2bulge9 is reported in fold induction of GFP levels relative to the expression levels in the absence of xanthosine feeding as described in Figure 3.2. Expression data for experiments performed with L2bulge1 exhibit similar induction profiles and levels (data not shown).

The second system explores the application of these ribozyme switches to the *in vivo* sensing of metabolite production to demonstrate that these switches provide a non-invasive mechanism through which to transmit molecular information from cells. Nucleoside phosphorylase activities resulting in *N*-riboside cleavage of purine nucleosides have been identified in various organisms[28]. We observe that feeding xanthosine to our yeast cultures results in the production of xanthine, a product synthesized through riboside cleavage of xanthosine. Relatively high xanthine accumulation was detected in cell extracts between 24-48 h after substrate feeding by HPLC analysis (Supplementary Figure 3.7). Xanthosine accumulation was detected in cell extracts at earlier times, indicating that specific levels of intracellular xanthosine accumulation may be required for efficient conversion to xanthine, possibly due to a high K_m value for this enzyme. The theophylline aptamer employed in our switch platforms possesses a reduced binding affinity for xanthine (27-fold lower than

theophylline)[23]. We employed two 'ON' switch constructs (L2bulge1 and L2bulge9) for the *in vivo* detection of xanthine production in cultures fed the precursor xanthosine (Figure 3.7A). GFP levels in cells fed xanthosine rose steadily between 24-40 h post-feeding in correlation with HPLC data (Figure 3.7B), illustrating the non-invasive metabolite-sensing capabilities of these switches through transmitting changes in metabolite accumulation to changes in reporter expression levels. This application demonstrates the scalability (DP1) of our switch platform, in which the unique properties of the sensing platform employed in this control system enable broad implementation in diverse applications not generally accessible by other regulatory systems.

3.3. Discussion

A key component in the development of an RNA-based framework for engineering ligand-controlled gene regulatory systems is captured within DP5: a mechanism through which to reliably transmit information between distinct domains of the molecule. The strand-displacement and helix-slipping mechanisms demonstrate different strengths and weaknesses as standardized means of transmitting information from the aptamer domain to the ribozyme domain. Only 7 out of the 26 tested communication modules exhibited regulatory activity in our system. In addition, all of the functional communication module sequences demonstrate 'OFF' activity in our *in vivo* system, whereas one of these sequences (cmd) exhibited 'ON' activity in an *in vitro* system[19]. These results indicate that *in vitro* functionality of these elements is selectively translated to *in vivo* activity due to their sensitivity to surrounding sequences. Furthermore, modularity studies performed on this platform indicate that the helix slipping mechanism is not amenable to modular domain swapping strategies. In contrast, we

have demonstrated that strand displacement exhibits greater reliability as an information transmission mechanism in our platform and is characterized by engineering properties such as modular assembly, rational *de novo* design, flexible induction and repression profiles, and response programmability. Although not preferred for the rational design strategies presented here, our helix slipping platform can be employed for the effective generation of new ribozyme switches by *in vivo* screening for helix slipping elements that function with new aptamer sequences, different regulatory ranges, and flexible regulatory profiles. In addition, screening strategies may represent a powerful alternative when rational design strategies fail. For example, we applied our rational design strategies to the construction of strand displacement-based ribozyme switches that modulate cleavage through stem I (L1bulge1-6 in Supplementary Table 3.3). Although these design strategies were successfully applied to stem II, they did not result in functional switches when applied to stem I. These results indicate that screening strategies may be more effective in generating ribozyme switches that modulate activity through stem I.

We have developed and demonstrated universal RNA-based regulatory platforms called ribozyme switches using engineering design principles. This work describes a framework for the reliable *de novo* construction of modular, portable, and scalable control systems that can be used to achieve flexible regulatory properties, such as up- and down-regulation of target expression levels and tuning of regulatory response to fit application-specific performance requirements, thereby expanding the utility of our platforms to a broader range of applications. For example, these switch platforms may be applied to the construction of transgenic regulatory control systems that are responsive to cell-permeable, exogenous molecules of interest for a given cellular network. In regulating sets of functional

proteins, these switches can act to rewire information flow through cellular networks and reprogram cellular behavior in response to changes in the cellular environment. In regulating reporter proteins, ribozyme switches can serve as synthetic cellular sensors for diverse input molecules to monitor temporal and spatial fluctuations in the levels of their target molecules. The switch platforms described here represent powerful tools for constructing ligand-controlled gene regulatory systems tailored to respond to specific effector molecules and enable regulation of target genes in various living systems, and due to their general applicability our platforms offer broad utility for applications in synthetic biology, biotechnology, and health and medicine.

3.4. Materials and Methods

3.4.1. Plasmid, switch construction, and cell strains

Using standard molecular biology techniques[29], a modular characterization plasmid, pRzS, harboring the yeast-enhanced green fluorescence protein (yEGFP)[30] under control of a GAL1-10 promoter, was constructed and employed as a universal vector for the characterization of all ribozyme switches. For the ribozyme switch-mediated growth studies, the *yegfp* gene was replaced with the *his5* gene[27]. The engineered ribozyme constructs were generated by PCR amplification using the appropriate oligonucleotide templates and primers. All oligonucleotides were synthesized by Integrated DNA Technologies. All engineered ribozyme constructs were cloned into two unique restriction sites, *Avr*II and *Xho*I, 3 nucleotides downstream of the yEGFP stop codon and upstream of an ADH1 terminator.

Cloned plasmids were transformed into an electrocompetent *Escherichia coli* strain, DH10B (Invitrogen) and all ribozyme constructs were confirmed by subsequent sequencing

(Laragen, Inc). Confirmed plasmid constructs were transformed into a *Saccharomyces cerevisiae* strain (W303 *MATα his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1*) using standard lithium acetate procedures[31].

3.4.2. 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. RNA sequences that are predicted to adopt at least two stable equilibrium conformations (ribozyme inactive and active) were constructed and examined for functional activity.

3.4.3. Ribozyme characterization assays

S. cerevisiae cells harboring the appropriate plasmids were grown in synthetic complete medium supplemented with an appropriate dropout solution and sugar (2% raffinose, 1% sucrose) overnight at 30°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°C. An appropriate volume of concentrated effector stock (to the appropriate final concentration of theophylline or tetracycline) dissolved in medium or an equivalent volume of the medium (no effector control) was added to the cultures at the time of back dilution. In addition, at this time an appropriate volume of galactose (2% final concentration) or an equivalent volume of water were added to the cultures for the induced and non-induced controls, respectively. For specificity assays, an appropriate volume of a concentrated caffeine or doxycycline stock (final concentrations of 1 mM and 250 μM, respectively) was added to a separate culture. 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 Safire fluorescent plate reader (Tecan) and/or on a Cell Lab Quanta SC flow cytometer (Beckman Coulter). For temporal response assays, cultures were grown as described above in the absence of the appropriate effector and fluorescence data were taken every 30 min. After 4 h growth, appropriate volumes of the concentrated effector stock or plain medium were added to the cultures and fluorescence was monitored for several hours thereafter.

3.4.4. Cell growth regulation assays

For liquid culture assays, *S. cerevisiae* cells carrying appropriate plasmids were back diluted and grown according to procedures described above with minor modifications. A competitive inhibitor of the *his5* gene product, 3-amino-triazole (3AT), was added to a final concentration of 5 mM to increase the threshold level of histidine required for cell growth. Cultures were grown in various theophylline concentrations and the growth of each sample was monitored over a 24 h period. The theophylline-regulated growth at 24 h is reported in terms of OD_{600} readings measured on the Tecan. For plate-based assays, 10 μL of the back diluted culture samples was streaked on plates containing 0 and 5 mM theophylline. A higher concentration of 3AT (25 mM) was used in the plate-based assays to optimize visual assessment of theophylline-regulated cell growth.

3.4.5. Metabolite sensing assays

S. cerevisiae cells carrying appropriate plasmids were back diluted and grown according to procedures described above with minor modifications. Cultures were grown in the absence and presence of xanthosine (250 μM final concentration). To account for inducer

depletion, galactose was added to the cultures at 8 h time intervals to a 2% final concentration. Fluorescence levels of the samples were monitored over a 48 h period according to procedures described above. For HPLC analysis, cell extracts were prepared after appropriate growth periods following xanthosine feeding by rapid freezing of cell cultures in liquid nitrogen in the form of beads. Frozen cell beads were subsequently lysated by grinding using a mortar and pestle followed by extraction with methanol. Intracellular metabolite levels were analyzed using an HPLC system integrated with a mass spectrometer (HPLC-MS) (Agilent Technologies), which enables confirmation of metabolite peaks based on their corresponding molecular weights.

3.4.6. Fluorescence quantification

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. A non-induced 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.

Similar to previous reports[17, 32], we report gene expression levels as 'fold', where 1 fold is defined as the reporter gene expression level of sTRSV relative to the background fluoresence level. Ligand-directed regulatory effects are reported as fold gene expression levels normalized to the levels in the absence of effector. All fluorescence data and mean ±s.d. are reported from at least three independent experiments.

3.4.7. Quantification of cellular transcript levels

Briefly, total RNA was extracted employing standard acid phenol extraction methods[33] followed by cDNA synthesis and PCR amplification. cDNA was synthesized using gene-specific primers (Supplementary Table 3.3) and Superscript III Reverse Transcriptase (Invitrogen) following manufacturer's instructions. Relative transcript levels were quantified from the cDNA samples by employing an appropriate primer set and the iQ SYBR Green Supermix (BioRAD) according to manufacturer's instructions on an iCycler iQ qRT-PCR machine (BioRAD). The resulting data were analyzed with the iCycler iQ software according to manufacturer's instructions. Transcript levels of switch constructs were normalized to that of the endogenous *actI* gene[34] using *actI*-specific primers.

3.5. Supplementary Information

Supplementary Text 3.1: *Glossary of terms*

actuator domain	A switch domain that encodes the system control function. As used here, the actuator domain encodes the gene regulatory function and is comprised of a hammerhead ribozyme sequence.

communication module	A sequence element that typically forms an imperfectly paired double-stranded stem that can adopt different base pairs between nucleotides through a 'slip-structure' mechanism. As used here, a communication module is a type of information transmission domain that transmits the binding state of the aptamer domain to the adjacent actuator domain through a helix slipping mechanism. As demonstrated in this work, a communication module does not act in a modular fashion with other switch domains. The term is retained here from earlier work in the field of nucleic acid engineering.
competing strand	The nucleic acid sequence within a strand displacement domain that is bound to the general transmission region of the switch when the sensor domain is in the restored conformation (i.e., in the presence of ligand). The competing strand competes for binding with the switching strand, which is initially bound to this transmission region in the absence of ligand.
component	A part of a system that encodes a distinct activity or function.
composability	A property of a system that indicates its ability to be comprised of components that can be selected and

	assembled in a modular fashion to achieve a desired system performance. As used here, composability refers to the ability of the individual domains of the control system to be modularly linked without disrupting their activities.
engineering design principle	A required property of a constructed system that enables use by others.
framework	A basic conceptual structure that is used to solve a complex product design issue. As used here, the framework is used to reliably design and construct specific instances of RNA switches. The conceptual structure of our framework is comprised of specified engineering design principles and design strategies that enable extensible and reusable system design.
helix slipping domain	A subset of information transmission domains that act through a helix slipping mechanism. The helix slipping domain is also referred to as the communication module.
helix slipping mechanism	An information transmission mechanism that is based on an information transmission domain that functions through a helix slipping event and does not allow for rational design. Such a helix slipping event utilizes a communication module (or helix slipping domain) within

	the general transmission region of the switch (the base stem of the aptamer) to result in disruption or restoration of the actuator domain in response to restoration of the sensor domain.
information transmission domain	A switch domain that encodes the function of transmitting information between the sensor domain and the actuator domain.
information transmission mechanism	A general mechanism for transmitting information between the sensor domain and the actuator domain of a switch. As used here, this mechanism regulates the activity of the actuator domain in response to the binding state of the sensor domain.
modular	A property of a system comprised of modules, which indicates that the modules can be interchanged as parts without changing the interface between modules or the modules themselves.
module	A self-contained system component that has a well-defined interface with other system components.
platform	A general framework on which specific applications can be implemented. As used here, the platform enables specific instances of switches to be built in a standardized

	manner.
portability	A property of a system that indicates its ability to be implemented in environments different from that which it was originally designed. As used here, portability refers to the ability of the control system to be implemented in different organisms.
reliability	A property of a system that indicates its ability to perform and maintain its functions under a set of specified conditions. As used here, reliability refers to the ability of the information transmission domain to standardize the transmission of information between the sensor and actuator domains.
scalability	A property of a system that indicates its ability to handle increasing work. As used here, scalability refers to the ability of the control system to be implemented across broad application space by being able to forward design its response to different molecular information.
switch	A molecule that can adopt at least two different conformational states, where each state is associated with a different activity of the molecule. Often a ligand can bind to one or more conformations of the switch, such that

	the presence of the ligand shifts the equilibrium distribution across the adoptable conformations and therefore regulates the activity of the switch molecule. As used here, switch refers to an RNA molecule that can adopt different structures that correspond to different gene-regulatory activities. An RNA switch is then a ligand-controlled gene regulatory system.
switch domain	A component of a switch that encodes a distinct activity or function.
switching strand	The nucleic acid sequence within a strand displacement domain that is bound to the general transmission region of the switch when the sensor domain is in the disrupted conformation (i.e., in the absence of ligand). The switching strand is displaced by the competing strand in the presence of ligand.
sensor domain	A switch domain that encodes a ligand-binding function. As used here, the sensor domain is comprised of an RNA aptamer sequence.
strand displacement domain	A subset of information transmission domains that act through a strand displacement mechanism.

strand displacement mechanism	An information transmission mechanism that is based on the rational design of an information transmission domain that functions through a strand displacement event. Such a strand displacement event utilizes competitive binding of two nucleic acid sequences (the competing strand and the switching strand) to a general transmission region of the switch (the base stem of the aptamer) to result in disruption or restoration of the actuator domain in response to restoration of the sensor domain.
universal	A system property that indicates its ability to maintain function across different applications, environments, and component interfaces. As used here, a universal system is composed of the five engineering design principles (scalability, portability, utility, composability, and reliability) and results in the specified extensible platform for RNA switch construction.
utility	A property of a system that indicates its ability to be of practical use. As used here, utility refers to the ability of the control system to interface with different functional level components to enable forward design of the function that is being controlled by the system.

Supplementary Text 3.2: *Ribozyme control constructs for loop sequence coupling and stem integration controls*

To establish and make useful our design strategy we constructed and characterized a series of ribozyme controls. We characterized the regulatory activity of our ribozyme constructs within a modular ribozyme characterization system in the eukaryotic model organism *Saccharomyces cerevisiae* (Figure 3.1A). First, an inactive ribozyme control (sTRSV Contl, Supplementary Figure 3.1A) was constructed to adopt the same structural motif as sTRSV (Figure 3.1A), while carrying a scrambled catalytic core sequence. Second, a synthetic sTRSV ribozyme (hhRz I) that contains closed loops in stems II and III and is embedded through stem I was constructed as a stem integration control (Supplementary Figure 3.1A). Finally, we constructed four loop sequence controls. In one set, stem loops I and II (L1R and L2R, respectively) were replaced by the theophylline aptamer TCT8-4[23] (Supplementary Figure 3.1B), and in another set, the theophylline aptamer was coupled directly to sequences in loops I and II (L1Theo and L2Theo, respectively) (Supplementary Figure 3.1C). sTRSV exhibits a 50-fold reduction in target expression levels relative to sTRSV Contl (Supplementary Figure 3.1D). HhRz I, L1R, and L2R exhibit similar target expression levels to that of sTRSV Contl, suggesting that ribozyme activity was abolished in these constructs. In contrast, L1Theo and L2Theo exhibit significantly lower target expression levels relative to sTRSV Contl. L1Theo and L2Theo were employed as the primary base constructs in engineering our synthetic ribozyme switch platforms. In addition, scrambled core versions of L1Theo and L2Theo exhibit no theophylline-dependent shifts in gene expression (data not shown), indicating that theophylline binding in that region of the transcript alone is not responsible for the observed regulatory effects. Taken together, we

find that our design strategy enables the construction of a universal ribozyme switch platform that satisfies the design principles of portability, utility, and composability.

Supplementary Text 3.3: *Rational tuning strategies for strand displacement-based switches*

A series of nine tuned 'ON' switches were constructed from L2bulge1 as a base structure by employing rational energetic tuning strategies developed in this work. This strategy is based on the effects of altering the predicted free energies of a particular conformation ($-\Delta G$) and the predicted difference between the free energies of two conformations ($\Delta\Delta G$) on RNA conformational dynamics, or the ability of the RNA molecule to distribute between these two conformational states. Supplementary Table 3.2 lists free energies ($-\Delta G$) of ribozyme active and inactive conformations and the energy difference ($\Delta\Delta G$) between the free energies of these two conformations. Specifically, lowering values for either of these energetic measurements ($-\Delta G$ or $\Delta\Delta G$) is expected to make it easier for a particular RNA molecule to switch between the conformational states in question. Therefore, there is an anticipated optimum conformational energy and energetic difference between conformations to achieve the desired range of switching in response to effector concentration (i.e., energy measurements too high will result in stable non-switch designs, and energy measurements or energy difference measurements too low will result in fairly equal distributions between the two conformational states and lower switching capabilities). It is also expected, then, that one can "push" switches into a non-switch state by moving away from this energetic optimum. This strategy was examined in a series of tuning experiments described below.

L2bulge2 and L2bulge3 (Supplementary Figure 3.5) replace canonical base pairs in the aptamer base stem of the ribozyme inactive conformation of L2bulge1 with U-G wobble pairs. As a result of these destabilizing alterations, both equilibrium conformations (ribozyme active and ribozyme inactive) become less thermodynamically stable than those of L2bulge1, as estimated from their predicted free energies (-ΔG). In addition, the energy required to switch between the two equilibrium conformations was maintained similar to that of L2bulge1, as estimated by the difference between the free energies of the two conformations (ΔΔG). Ribozyme assays indicate that both L2bulge2 and L2bulge3 exhibit smaller dynamic ranges than that of L2bulge1 (Figure 3.4C and Supplementary Figure 3.8). It is proposed that the lower stabilities of the conformational states enable more frequent dynamic switching between the two equilibrium conformations and therefore lower the difference in distribution favoring one state over the other.

L2bulge4 (Supplementary Figure 3.5) incorporates an additional G-U wobble pair within the aptamer base stem of the ribozyme inactive conformation of L2bulge1. However, this aptamer stem extension does not result in an appreciable predicted change in the thermodynamic stabilities of the equilibrium conformations or the energy required to switch between the two equilibrium conformations when compared to L2bulge1. Ribozyme assays indicate that L2bulge4 exhibits a dynamic range in response to theophylline levels similar to that of L2bulge1 (Figure 3.4C and Supplementary Figure 3.8).

L2bulge5 (Supplementary Figure 3.5) incorporates an additional canonical base pair (A-U) within the aptamer base stem of L2bulge1. As a result of this stabilizing alteration, the conformation of the ribozyme switch, in which the aptamer structure is formed and the catalytic core is disrupted (ribozyme inactive), is increased in stability and as stable as the

conformation in which the catalytic core is not disrupted (ribozyme active). The increased stability of the ribozyme inactive conformation in L2bulge5 in comparison to L2bulge1 and L2bulge4 indicates that the equilibrium distribution between these two conformations will shift to favor the ribozyme inactive conformation. Ribozyme assays indicate that L2bulge5 exhibits significantly higher GFP expression levels in the absence and presence of theophylline compared to those of L2bulge1 and L2bulge4, such that the theophylline-regulated increase in gene expression is similar to that of L2bulge3 but different in regulatory dynamic ranges (Figure 3.4C and Supplementary Figure 3.8).

Two switches in this series, L2bulge6 and L2bulge7, were constructed to demonstrate the ability of this tuning strategy to "push" the ribozyme switch constructs out of a switchable energetic range and approach non-switching extremes. L2bulge6 (Supplementary Figure 3.5) was designed to energetically favor the conformation, in which the aptamer structure is formed and the catalytic core is disrupted, (ribozyme inactive) in the absence of theophylline by introducing a stabilizing G-C base pair into the aptamer stem of this conformation. Since the aptamer conformation is expected to be favored in L2bulge6, the presence of theophylline is expected to have little or no effect on the conformational dynamics of this switch. L2bulge7 (Supplementary Figure 3.5) was designed to energetically favor the conformation, in which the aptamer structure is not formed and the catalytic core is undisrupted (ribozyme active), by introducing a stabilizing U-A base pair into the stem extending from loop II in this conformation. As the stability of the ribozyme active conformation is significantly higher than that of the ribozyme inactive conformation for L2bulge7, the presence of theophylline is expected to have little effect on the conformational dynamics of this ribozyme switch. Ribozyme assays indicate that L2bulge7 exhibits very low

GFP expression levels and L2bulge6 exhibits very high GFP expression levels in the presence and absence of theophylline (Supplementary Figure 3.8). As rationally designed, both constructs exhibit little increase in target expression levels in response to theophylline by energetically favoring one of the two conformational states (Figure 3.4C).

L2bulge 8 (Supplementary Figure 3.5) was modified from L2bulge7 by replacing the canonical base pair (U-A) with a wobble base pair (U-G), thereby reducing the stability of the ribozyme active conformation of L2bulge7 and allowing it to adopt the ribozyme inactive conformation. Similarly, L2bulge 9 (Supplementary Figure 3.5) was modified in such a way to reduce the energy difference between the two conformations of L2bulge7. Ribozyme assays indicate that L2bulge8 and L2bulge9 exhibit theophylline-dependent up-regulation of target gene expression in accordance with the reduced stabilities of the ribozyme active conformations and energy differences between the two adoptable conformations for each of these switch constructs (Figure 3.4C and Supplementary Figure 3.8).

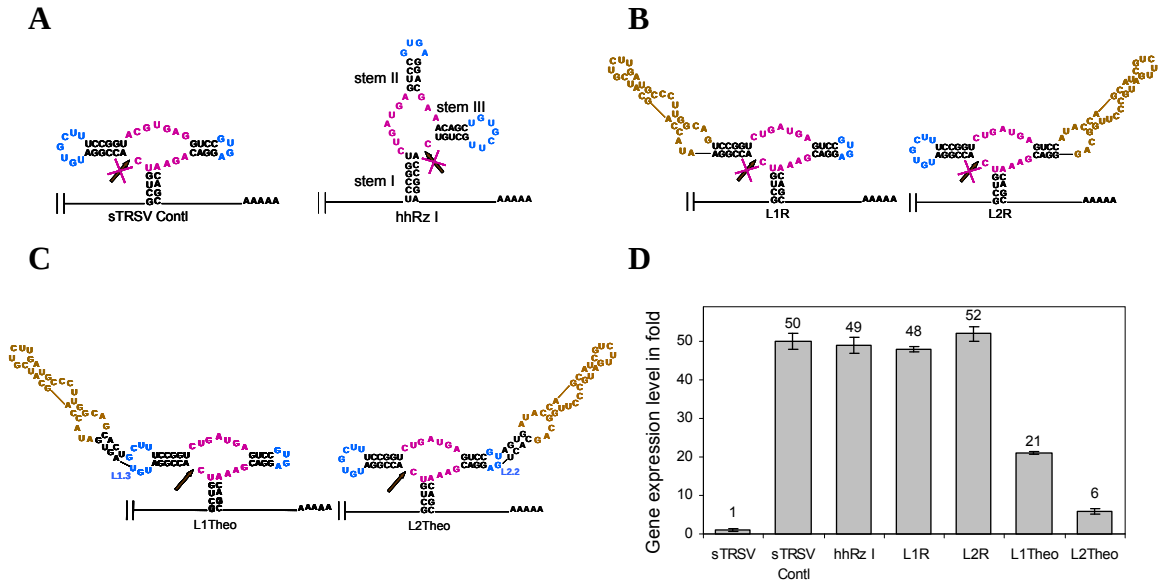
In addition, a series of three tuned ‘OFF’ switches were constructed by using rational energetic tuning strategies from L2bulgeOff1 as a base structure. L2bulgeOff2 and L2bulgeOff3 were constructed to demonstrate tunability of the ‘OFF’ switch platform using similar energetic design strategies (Supplementary Figure 3.5). These switch variants exhibit different theophylline-responsive dynamic ranges from that of L2bulgeOff1 (Figure 3.4D and Supplementary Figure 3.8).

Flow cytometry analysis of the tuned ribozyme switch series demonstrate that the tuned switches exhibit corresponding shifts in the mean fluorescence of the cell populations in the presence and absence of theophylline (Supplementary Figure 3.6). The relative dynamic ranges of the switches across the full regulatory range bracketed by the ribozyme

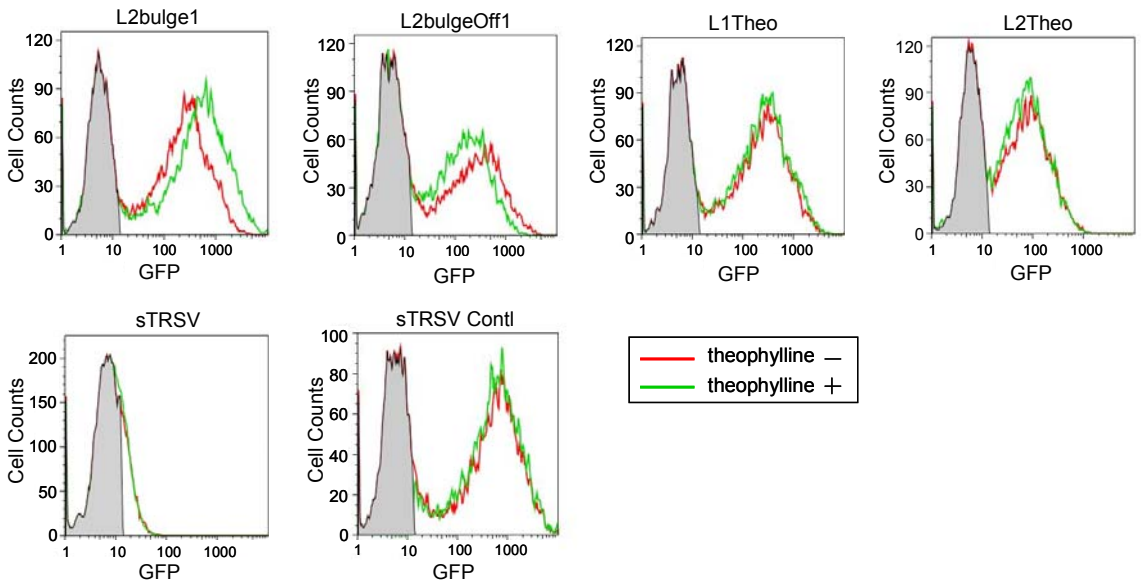
active and inactive controls, sTRSV and sTRSV Contl respectively, are presented in Supplementary Figure 3.8.

Among the twelve tuned switches (both 'ON' and 'OFF'), the dynamic regulatory ranges of most of these switches are in agreement with our rational tuning strategies based on the -ΔG and ΔΔG values predicted by RNAstructure 4.2. Two exceptions are noted: L2bulge9 and L2bulgeOff3. L2bulge9 exhibits a larger dynamic regulatory effect despite its higher ΔΔG than L2bulge8. L2bulgeOff3 exhibits a smaller dynamic regulatory effect despite its smaller ΔΔG than L2bulgeOff2. However, it is more difficult to make a direct comparison between L2bulgeOff2 and L2bulgeOff3, as both conformations of L2bulgeOff3 are significantly more stable than those of L2bulgeOff2, likely resulting in L2BulgeOff3 less frequently switching between its two conformations and thus enabling this molecule to get 'trapped' in its lower free energy states. In addition, outliers may also arise because the RNAstructure program predicts these energy values based on the secondary structure of a particular conformation and does not take into consideration energy contributions from tertiary interactions (that have been observed in prior work[16]) in estimating these energies. Nevertheless, we demonstrate that energetic predictions based solely on secondary structure are useful for our rational tuning design strategies. The different dynamic regulatory ranges exhibited by our tuned switches in response to their specific effector (Supplementary Figure 3.8) validate that such response programming can be achieved by altering the nucleotide composition of the information transmission region within a switch, thereby demonstrating the interdependence between RNA sequence, structure, and function.

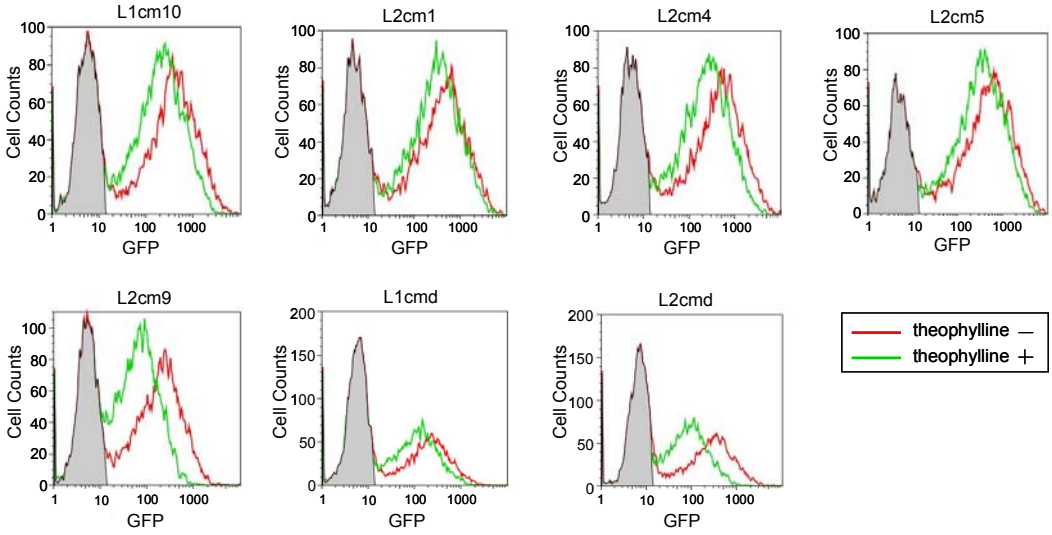
Supplementary Figures and Tables



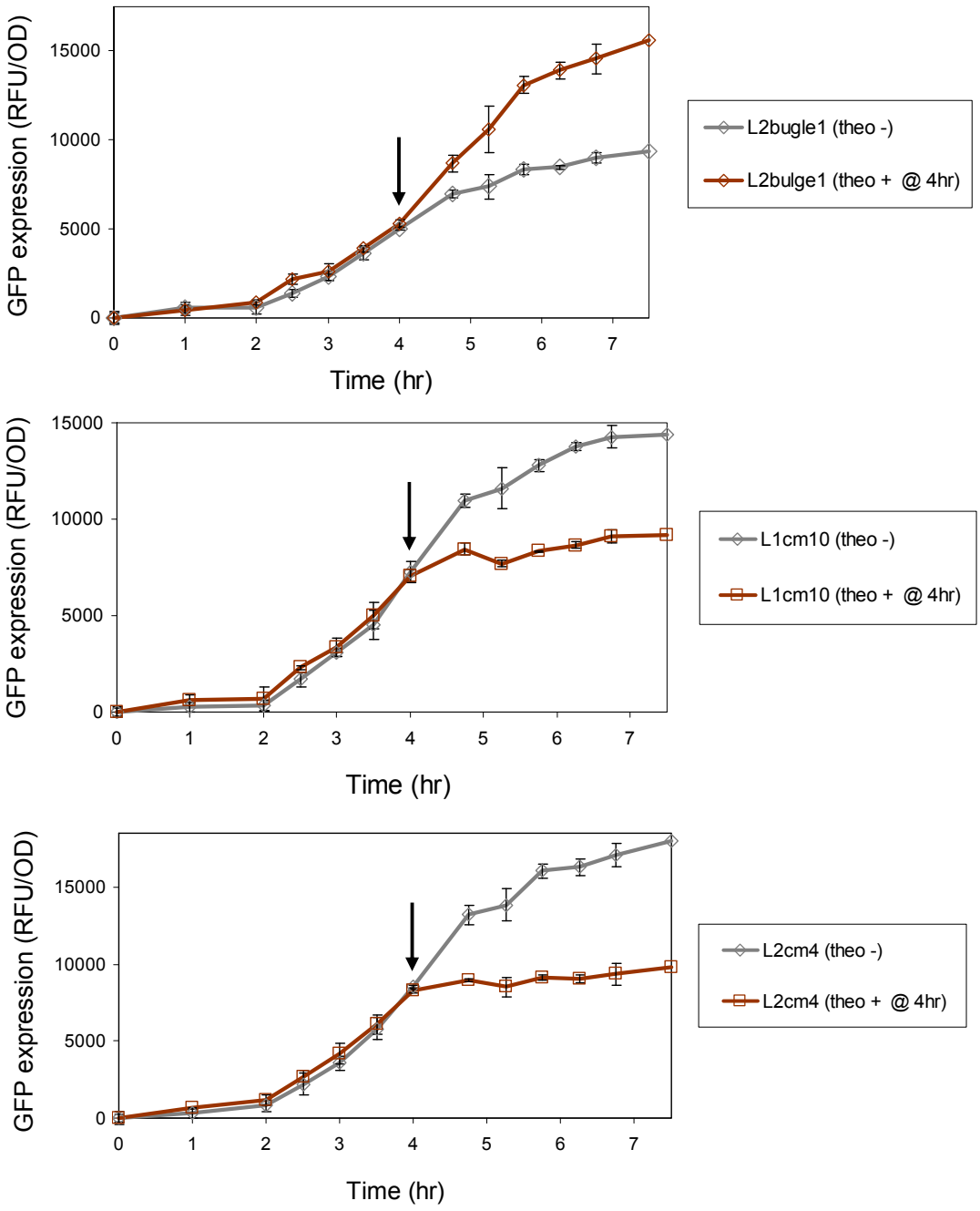
Supplementary Figure 3.1. Control constructs supporting the design strategy for engineering ligand-regulated ribozyme switches. Color schemes are as follows: catalytic core, purple; aptamer sequences, brown; loop sequences, blue; brown arrow, cleavage site. (A) Sequences of the ribozyme (sTRSV Contl) and stem integration (hhRz I) controls. (B) Sequences of the loop sequence controls in which the loop I and II sequences are replaced by the theophylline aptamer (L1R and L2R, respectively). (C) Sequences of the loop sequence controls in which the theophylline aptamer is connected directly to the loop I nucleotides through L1.3 and L1.4 (L1Theo) and the loop II nucleotides through L2.2 and L2.3 (L2Theo). (D) Gene expression levels (in fold) of the control constructs. 1-fold is defined as the reporter gene expression level of sTRSV relative to that of the background fluorescence level. The mean ±s.d. from at least three independent experiments is shown.



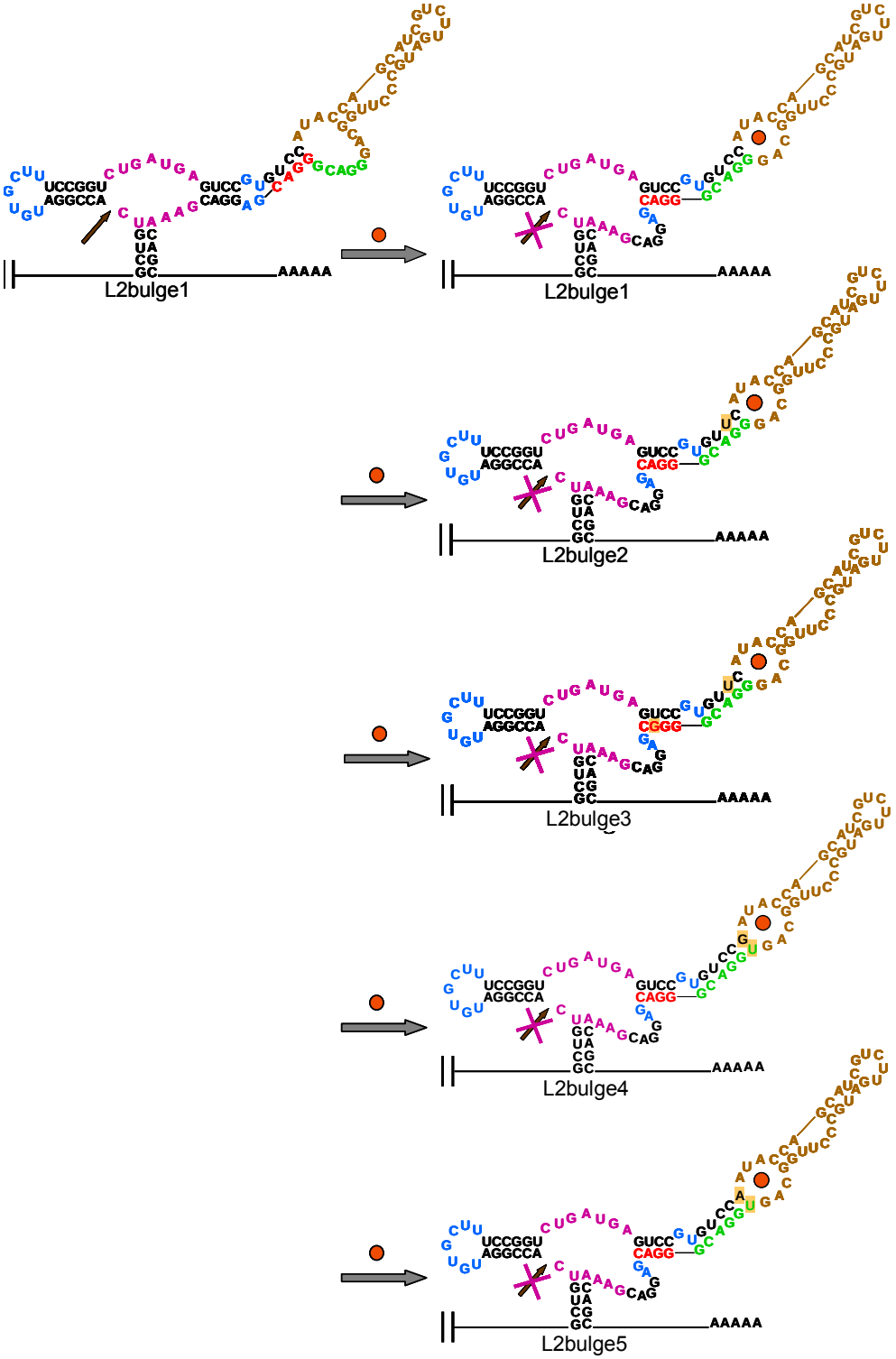
Supplementary Figure 3.2. Flow cytometry histograms of L2bulge1, L2bulgeOff1, and the ribozyme control cell populations grown in the presence (+) and absence (-) of 5 mM theophylline. Red line: cell populations grown in the absence of theophylline; green line: cell populations grown in 5 mM theophylline; shaded population: cell populations indicative of the non-induced cell population, shaded here to indicate the portion of cells in the population that have lost the plasmid and exhibit non-induced, or background, levels of autofluorescence. Histograms are representative of three independent experiments.

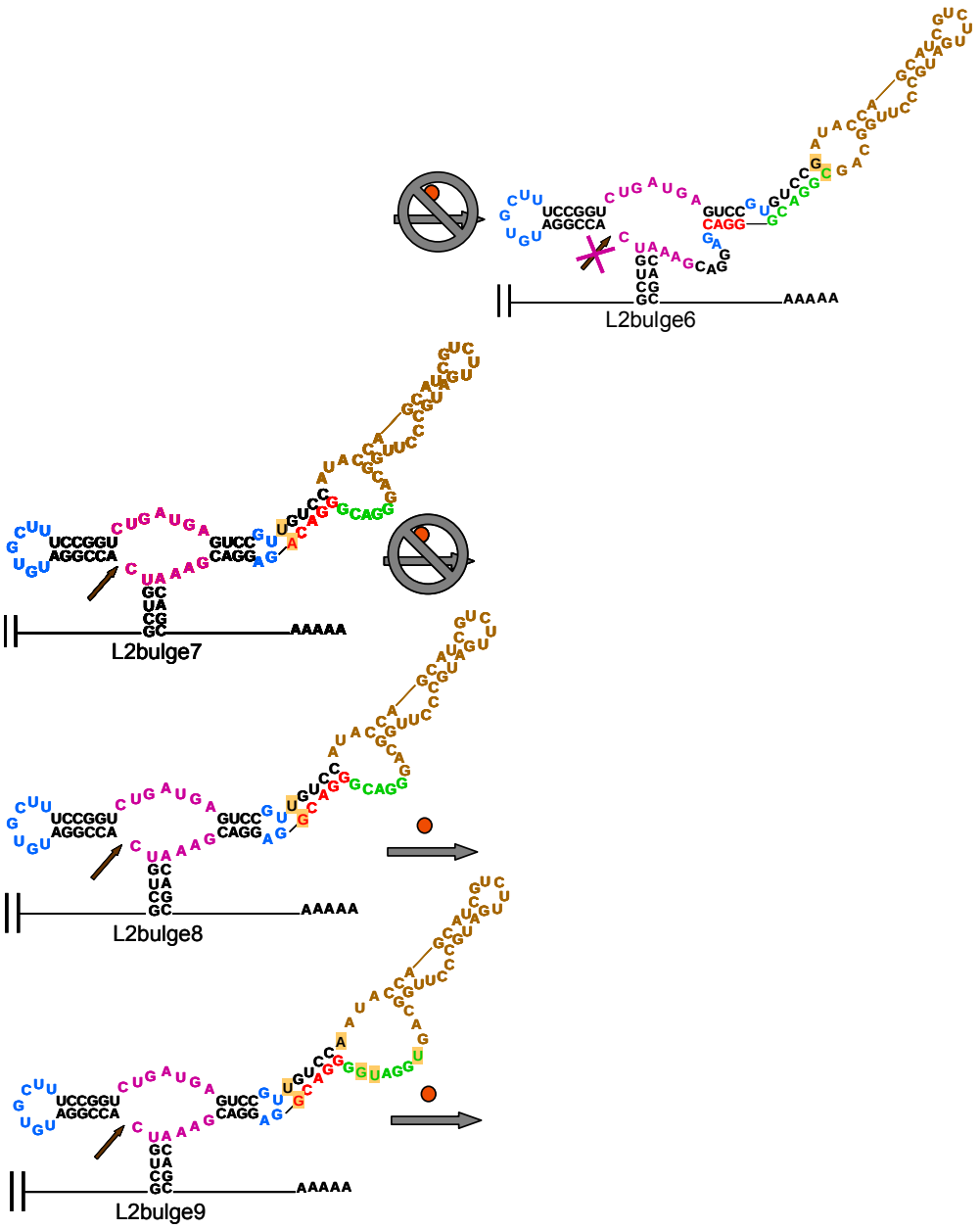


Supplementary Figure 3.3. Flow cytometry histograms of the helix slipping-based ribozyme switch cell populations grown in the presence (+) and absence (-) of 5 mM theophylline. Population data is measured and reported as described in Supplementary Figure 3.2. Histograms are representative of three independent experiments.

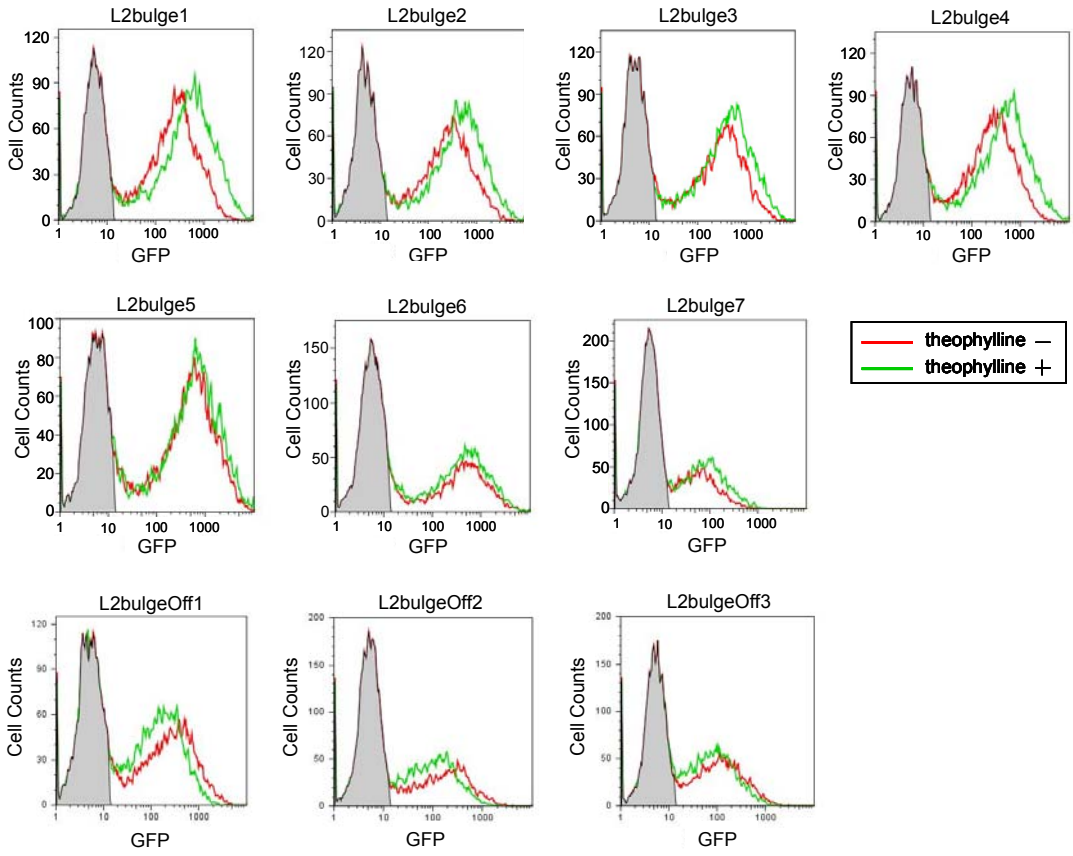


Supplementary Figure 3.4. Temporal responses of L2bulge1, L1cm10, and L2cm4 in response to the addition of 5 mM theophylline (final concentration). The time point at which theophylline was added to the cultures is indicated by an arrow. Brown: 5 mM theophylline added to growing cultures; gray: no theophylline added to growing cultures. Gene expression levels are reported as RFU/OD by dividing fluorescence units by the OD600 of the cell sample and subtracting the background fluorescence level. L2bulge1 exhibits up-regulation of GFP levels in response to the addition of theophylline; L1cm10 and L2cm4 exhibit down-regulation of GFP levels in response to theophylline addition. The mean ± s.d. from at least three independent experiments is shown for all graphs.

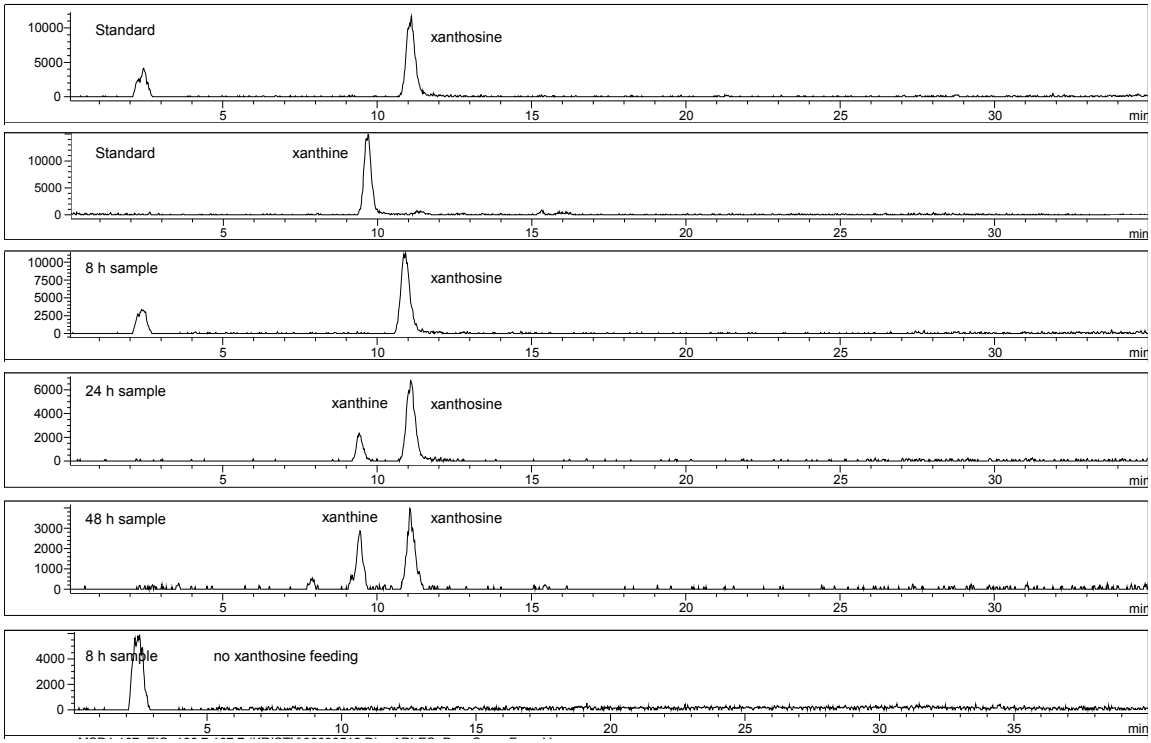

L2bulge1
L2bulge1
L2bulge2
L2bulge3
L2bulge4
L2bulge5



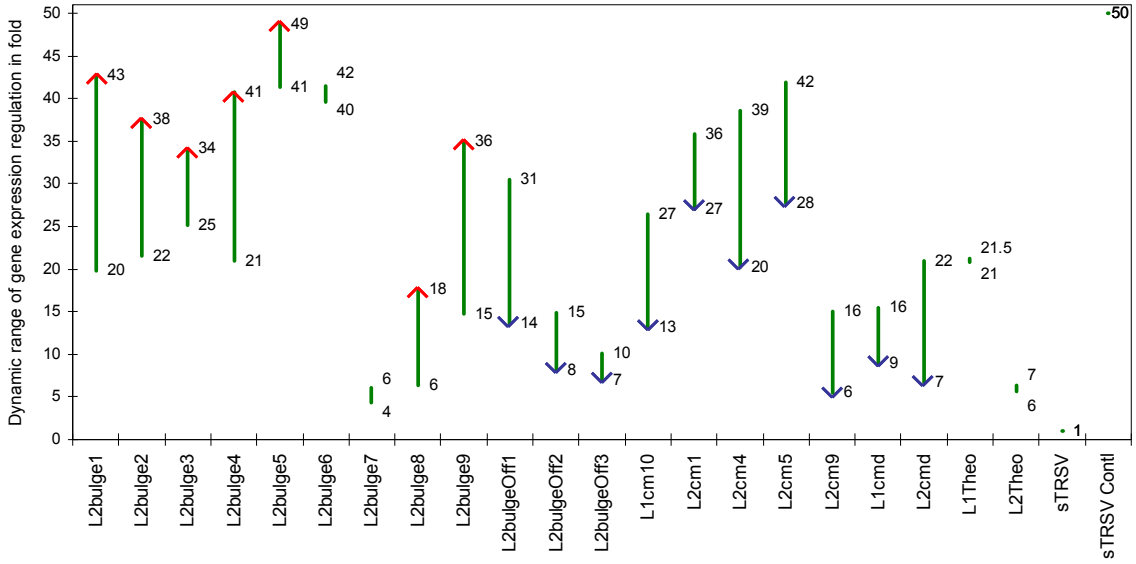
Supplementary Figure 3.5. Sequences and structures of tuned ribozyme switches in the L2bulge series. The nucleotides altered from the parent construct, L2bulge1, are highlighted. The two stable equilibrium conformations, ribozyme active and inactive conformations, are indicated for the parent ribozyme switch. The ribozyme active conformations of L2bulge2-5 are not shown as they are similar to L2bulge1. L2bulge6 and L2bulge7 assume a single predominant conformation, ribozyme inactive and ribozyme active, respectively, and do not undergo theophylline-induced conformational switching. L2bulge8 and L2bulge9, modified from L2bulge7 by reducing the stability of the ribozyme active conformation and the energy difference between the two conformations of L2bulge7, now become capable of switching. For these two modified switch constructs, only the ribozyme active conformations are shown, as their ribozyme inactive conformations are similar to those of the other switches illustrated.



Supplementary Figure 3.6. Flow cytometry histograms of the tuned ribozyme switch series cell populations grown in the presence (+) and absence (-) of 5 mM theophylline. Population data is measured and reported as described in Supplementary Figure 3.2. Histograms are representative of three independent experiments.



Supplementary Figure 3.7. Detection of intracellular accumulation of the substrate xanthosine and the product xanthine over three different time points. Accumulation of xanthosine is observed at earlier time points. Conversion of xanthosine to xanthine was detected at 24 h after substrate feeding and a higher accumulation of xanthine was detected at 48 h after substrate feeding.



Supplementary Figure 3.8. Dynamic ranges of regulation of the ribozyme switches and controls engineered in this work. The regulatory effects at 5 mM theophylline are reported on a full transcriptional range spectrum scale without normalization to the corresponding base expression level of each switch in the absence of effector (0 mM). Little or no effector-mediated gene regulatory effect is observed in the non-switch control constructs. Gene expression fold is defined as previously where 1 fold is equivalent to the reporter gene expression level of sTRSV relative to the background fluorescence level. sTRSV is the most active ribozyme construct exhibiting the lowest gene expression level and sTRSV Contl is the most inactive ribozyme construct exhibiting the highest gene expression level, providing a 50-fold range as the full spectrum equivalent to a total of 50 folds. Arrows indicate the direction of regulation as an increasing concentration of theophylline. These switches offer diverse dynamic ranges of regulation and thus provide a broader utility to fit specific applications of interest. Data are reported from three independent experiments and the ± s.d. is the same as that reported in the manuscript figures.

Supplementary Table 3.1. Relative steady-state ribozyme switch and ribozyme control transcript levels in the presence and absence of theophylline.

constructs	0 mM theopylline	5 mM theophylline	regulatory effect
sTRSV	0.08±0.01	0.11±0.01	little
sTRSV Contl	1.00±0.06	1.10±0.04	little
L2bulge1	0.49±0.04	0.77±0.10	up-regulation
L1cm10	0.66±0.05	0.43±0.06	down-regulation
L2cm4	0.67±0.05	0.38±0.06	down-regulation

Quantitative RT-PCR (qRT-PCR) analysis was performed on L2bulge1, L1cm10, L2cm4, satellite RNA of tobacco ringspot virus (sTRSV), and sTRSV control (sTRSV Contl). Transcript levels in the presence or absence of theophylline are reported as fractions relative to those of sTRSV Contl. L2bulge1 exhibits a higher steady-state level of target transcript, while L1cm10 and L2cm4 exhibited lower steady-state target transcript levels in the presence of 5 mM theophylline than in the absence of theophylline. The ribozyme controls, sTRSV and sTRSV Contl, exhibited little effect on steady-state transcript levels due to the presence of theophylline. In addition, relative steady-state levels of these switches corresponded to the relative GFP expression levels as determined through the functional ribozyme switch assays. All data are reported from three independent experiments.

Supplementary Table 3.2. Free energies (-ΔG, kcal/mol) of individual conformations (ribozyme-active and -inactive) and the energy difference (ΔΔG, kcal/mol) between the free energies of these two conformations predicted by RNAstructure 4.2.

switch constructs	free energy of aptamer-unbound conformation (-ΔG)	free energy of aptamer-bound conformation (-ΔG)	free energy difference (ΔΔG)
ON switches	ribozyme active	ribozyme inactive	
L2bulge1	38.9	38.1	0.8
L2bulge2	36.0	35.2	0.8
L2bulge3	35.5	34.6	0.9
L2bulge4	39.5	38.8	0.7
L2bulge5	39.5	39.5	0.0
L2bulge6	39.2	40.5	-1.3
L2bulge7	40.2	36.5	3.7
L2bulge8	39.4	38.0	1.4
L2bulge9	39.3	37.7	1.6
OFF switches	ribozyme inactive	ribozyme active	
L2bulgeOff1	39.3	38.6	0.7
L2bulgeOff2	39.3	37.2	2.1
L2bulgeOff3	39.9	38.2	1.7

Supplementary Table 3.3. Primer and additional ribozyme construct sequences.

name	oligonucleotide sequences	comments
5' spacer	5' AAACAAACAAA	spacer sequence preceding each construct
3' spacer	5' AAAAAGAAAAATAAAAA	spacer sequence following each construct
L1-2 fwd	5' GACCTAGGAAACAAACAAAGCTGTCACC	forward primer for all constructs
L1-2 rev	5' GGCTCGAGTTTTTATTTTTCTTTTTGCTGTTTCG	reverse primer for all constructs
yEGFP fwd	5' CGGTGAAGGTGAAGGTGATGCTACT	forward primer specific for qRT-PCR of *yegfp*
yEGFP rev	5' GCTCTGGTCTTGTAGTTACCGTCATCTTTG	reverse primer specific for cDNA synthesis and qRT-PCR of *yegfp*
ActI fwd	5' CGGTGAAGGTGAAGGTGATGCTACT	forward primer specific for qRT-PCR of *actI*
ActI rev	5' GCTCTGGTCTTGTAGTTACCGTCATCTTTG	reverse primer specific for cDNA synthesis and qRT-PCR of *actI*
L1bulge1	5' GCTGTCACCGGATGTACCGGAATACCAGCATCGTCTTGAT GCCCTTGGCAGTCTGGTCCGGTGCTTTCCGGTCTGATGAGT CCGTGAGGACGAAACAGC	no functional activity observed
L1bulge2	5' GCTGTCACCGGATGTACCGGAATACCAGCATCGTCTTGAT GCCCTTGGCAGTCCGGTCCGGTGCTTTCCGGTCTGATGAGT CCGTGAGGACGAAACAGC	no functional activity observed
L1bulge3	5' GCTGTCACCGGATGTACCGGAATACCAGCATCGTCTTGAT GCCCTTGGCAGTCTGGGTCCGGTGCTTTCCGGTCTGATGAGT CCGTGAGGACGAAACAGC	no functional activity observed
L1bulge4	5' GCTGTCACCGGATGTACCGGAATACCAGCATCGTCTTGAT GCCCTTGGCAGTCCGGGTCCGGTGCTTTCCGGTCTGATGAGT CCGTGAGGACGAAACAGC	no functional activity observed
L1bulge5	5' GCTGTCACCGGATGTACCGGAATACCAGCATCGTCTTGAT GCCCTTGGCAGTCTGGTTCCGGTGCTTTCCGGTCTGATGAGT CCGTGAGGACGAAACAGC	no functional activity observed
L1bulge6	5' GCTGTCACCGGATGTACCGGAATACCAGCATCGTCTTGAT GCCCTTGGCAGTCTGGATCCGGTGCTTTCCGGTCTGATGAGT CCGTGAGGACGAAACAGC	no functional activity observed
L1cm10	5'CCTAGGAAACAAACAAAGCTGTCACCGGATGTGCTTTCCGG TCTGATGAGTCCGTAAATGATACCAGCATCGTCTTGATGCCCTT GGCAGCTGCGAGGACGAAACAGCAAAAAGAAAAATAAAAA CTCGAG	This sequence represents a template for other communication module constructs through L1 by replacing the colored sequences with corresponding modules
L2cm4	5'CCTAGGAAACAAACAAAGCTGTCACCGGATGTCCTGGATACC AGCATCGTCTTGATGCCCTTGGCAGTCATAGCTTTCCGGTCTG ATGAGTCCGTGAGGACGAAACAGCAAAAAGAAAAATAAAAA CTCGAG	This sequence represents a template for other communication module constructs through L2 by replacing the colored sequences with corresponding modules
cma	5'CCTT 5'ACGT	ref. 19, "class I induction element"
cm2	5'CCAGG 5'TTTGA	ref. 20
cm3	5'TCTGG 5'TCTTA	ref. 20
cm6	5'GGATG 5'CAAT	ref. 21
cm7	5'GGAGG 5'CCTT	ref. 21
cm8	5'ATACG 5'CGGT	ref. 21
cm11	5'TCGAG 5'CTCTA	ref. 21
cm12	5'AGGG 5'CTCTA	ref. 21

Acknowledgements

We thank K. Hawkins for assistance in controls and HPLC experiments and data analysis; A. Babiskin for supply of pRzS and assistance in qRT-PCR assays; K. Dusinberre, J. Michener, J. Liang for assistance in controls; E. Kelsic for assistance in image presentation; Y. Chen, K. Hoff for critical reading of the manuscript. 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. Endy, D. Foundations for engineering biology. Nature 438, 449-53 (2005).
2. Voigt, C. A. Genetic parts to program bacteria. Curr Opin Biotechnol 17, 548-57 (2006).
3. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci U S A 89, 5547-51 (1992).
4. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Res 25, 1203-10 (1997).
5. Mandal, M. & Breaker, R. R. Gene regulation by riboswitches. Nat Rev Mol Cell Biol 5, 451-63 (2004).
6. Kim, D. S., Gusti, V., Pillai, S. G. & Gaur, R. K. An artificial riboswitch for controlling pre-mRNA splicing. Rna 11, 1667-77 (2005).
7. An, C. I., Trinh, V. B. & Yokobayashi, Y. Artificial control of gene expression in mammalian cells by modulating RNA interference through aptamer-small molecule interaction. Rna 12, 710-6 (2006).
8. Bayer, T. S. & Smolke, C. D. Programmable ligand-controlled riboregulators of eukaryotic gene expression. Nat Biotechnol 23, 337-43 (2005).
9. Isaacs, F. J., Dwyer, D. J. & Collins, J. J. RNA synthetic biology. Nat Biotechnol 24, 545-54 (2006).

10. Bunka, D. H. & Stockley, P. G. Aptamers come of age - at last. Nat Rev Microbiol 4, 588-96 (2006).

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

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

13. Hermann, T. & Patel, D. J. Adaptive recognition by nucleic acid aptamers. Science 287, 820-5 (2000).

14. Birikh, K. R., Heaton, P. A. & Eckstein, F. The structure, function and application of the hammerhead ribozyme. Eur J Biochem 245, 1-16 (1997).

15. Marschall, P., Thomson, J. B. & Eckstein, F. Inhibition of gene expression with ribozymes. Cell Mol Neurobiol 14, 523-38 (1994).

16. 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).

17. Yen, L. et al. Exogenous control of mammalian gene expression through modulation of RNA self-cleavage. Nature 431, 471-6 (2004).

18. Koizumi, M., Soukup, G. A., Kerr, J. N. & Breaker, R. R. Allosteric selection of ribozymes that respond to the second messengers cGMP and cAMP. Nat Struct Biol 6, 1062-71 (1999).

19. Soukup, G. A. & Breaker, R. R. Engineering precision RNA molecular switches. Proc Natl Acad Sci U S A 96, 3584-9 (1999).

20. Soukup, G. A., Emilsson, G. A. & Breaker, R. R. Altering molecular recognition of RNA aptamers by allosteric selection. J Mol Biol 298, 623-32 (2000).

21. Kertsburg, A. & Soukup, G. A. A versatile communication module for controlling RNA folding and catalysis. Nucleic Acids Res 30, 4599-606 (2002).

22. Pelletier, J. & Sonenberg, N. Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency. Cell 40, 515-26 (1985).

23. Jenison, R. D., Gill, S. C., Pardi, A. & Polisky, B. High-resolution molecular discrimination by RNA. Science 263, 1425-9 (1994).

24. Berens, C., Thain, A. & Schroeder, R. A tetracycline-binding RNA aptamer. Bioorg Med Chem 9, 2549-56 (2001).

25. Hanson, S., Berthelot, K., Fink, B., McCarthy, J. E. & Suess, B. Tetracycline-aptamer-mediated translational regulation in yeast. Mol Microbiol 49, 1627-37 (2003).

26. Koch, A. L. The metabolism of methylpurines by Escherichia coli. I. Tracer studies. J Biol Chem 219, 181-8 (1956).

27. Nishiwaki, K. et al. Structure of the yeast HIS5 gene responsive to general control of amino acid biosynthesis. Mol Gen Genet 208, 159-67 (1987).

28. Ogawa, J. et al. Purification, characterization, and gene cloning of purine nucleosidase from Ochrobactrum anthropi. Appl Environ Microbiol 67, 1783-7 (2001).

29. Sambrook, J. & Russell, D. W. Molecular cloning: a laboratory manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001).

30. Mateus, C. & Avery, S. V. Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry. Yeast 16, 1313-23 (2000).

31. 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).

32. Isaacs, F. J. et al. Engineered riboregulators enable post-transcriptional control of gene expression. Nat Biotechnol 22, 841-7 (2004).

33. Caponigro, G., Muhlrad, D. & Parker, R. A small segment of the MAT alpha 1 transcript promotes mRNA decay in Saccharomyces cerevisiae: a stimulatory role for rare codons. Mol Cell Biol 13, 5141-8 (1993).

34. Ng, R. & Abelson, J. Isolation and sequence of the gene for actin in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 77, 3912-6 (1980).