

Cell-Targeted Regulation of
Gene Expression through
Synthetic RNA Devices

Thesis by
James Vincent Vowles

In Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy



CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2014

(Defended May 12, 2014)

© 2014

James Vincent Vowles

All Rights Reserved

Acknowledgements

This thesis would not have been possible without the unending support of Professor Christina Smolke. She has created a laboratory that nurtures young scientists while producing research of the highest caliber. Her insights have guided me through countless struggles and helped me develop as a scientist.

My deep appreciation goes out to all the past and present members of the Smolke laboratory. Every one of them has helped me in some way, large or small. Their breadth of expertise has been immensely valuable, so that when I am tackling a new challenge I can benefit from their greater skills and knowledge rather than going it alone. I am especially indebted to Maung Win, whose initial development of the ribozyme switch platform helped elevate the lab to the forefront of synthetic biology and provided me with a starting point for my project. For my entire graduate career I have been following the trail that he blazed. I also owe much of my success to Drew Kennedy, my partner on this project and my friend for nearly seven years. Drew has challenged me to be a better scientist, pointing out the holes in my reasoning and helping me to see the way forward from stubborn problems.

My thanks go to the members of my thesis committee, Professors Peter Dervan, Doug Rees, and Carl Parker. They offered fresh perspectives on my research and their support helped me leave behind failed projects to focus on more promising pursuits. Thanks also to Professor Harry Gray and Craig Wiggernhorn for facilitating the videoconferences that allowed me to meet with my committee without having to travel from Stanford to Caltech.

Finally, I must thank my friends and family. They supported me through grad school's most trying times, accepting my greatly decreased availability when my research demanded it. No one experienced this more strongly than my partner Ashley, who has often felt at war with the lab for my attention. She is the most loving and understanding person I have ever known and her support has been immeasurably valuable.

Abstract

The ability to interface with and program cellular function remains a challenging research frontier in biotechnology. Although the emerging field of synthetic biology has recently generated a variety of gene-regulatory strategies based on synthetic RNA molecules, few strategies exist through which to control such regulatory effects in response to specific exogenous or endogenous molecular signals. Here, we present the development of an engineered RNA-based device platform to detect and act on endogenous protein signals, linking these signals to the regulation of genes and thus cellular function.

We describe efforts to develop an RNA-based device framework for regulating endogenous genes in human cells. Previously developed RNA control devices have demonstrated programmable ligand-responsive genetic regulation in diverse cell types, and we attempted to adapt this class of cis-acting control elements to function in trans. We divided the device into two strands that reconstitute activity upon hybridization. Device function was optimized using an *in vivo* model system, and we found that device sequence is not as flexible as previously reported. After verifying the *in vitro* activity of our optimized design, we attempted to establish gene regulation in a human cell line using additional elements to direct device stability, structure, and localization. The significant limitations of our platform prevented endogenous gene regulation.

We next describe the development of a protein-responsive RNA-based regulatory platform. Employing various design strategies, we demonstrated functional devices that both up- and downregulate gene expression in response to a heterologous protein in a human cell line. The activity of our platform exceeded that of a similar, small-molecule-

responsive platform. We demonstrated the ability of our devices to respond to both cytoplasmic- and nuclear-localized protein, providing insight into the mechanism of action and distinguishing our platform from previously described devices with more restrictive ligand localization requirements. Finally, we demonstrated the versatility of our device platform by developing a regulatory device that responds to an endogenous signaling protein.

The foundational tool we present here possesses unique advantages over previously described RNA-based gene-regulatory platforms. This genetically encoded technology may find future applications in the development of more effective diagnostic tools and targeted molecular therapy strategies.

TABLE OF CONTENTS

Acknowledgements.....	iii
Abstract.....	v
Table of Contents.....	vii
List of Figures.....	ix
CHAPTER 1: Introduction.....	I-1
RNA is a versatile regulatory biomolecule.....	I-2
Engineered RNA devices in eukaryotes enable dynamic modulation of gene expression in response to molecular and environmental signals	I-3
Ligand-responsive ribozyme switches.....	I-6
Applications	I-11
Scope of thesis	I-12
References.....	I-14
CHAPTER 2: Development of an RNA device framework that targets endogenous genes in human cells	II-1
Abstract.....	II-2
Introduction.....	II-3
Results.....	II-8
Design of a trans-ribozyme-based regulatory element in human cells	II-8
Characterization of initial trans-ribozyme designs in a human cell line	II-11
Cis-ribozymes as a model for optimizing <i>in vivo</i> cleavage activity.....	II-15
Characterization of Type I cis-ribozymes in a human cell line.....	II-17
Development of an improved trans-ribozyme.....	II-18
Incorporation of ancillary elements in the trans-ribozyme transcript to direct stability, structure, processing, and localization	II-20
Discussion	II-23
Methods.....	II-27
Acknowledgements.....	II-36
References.....	II-37

Supplementary Tables.....	II-41
CHAPTER 3: Development of an RNA device framework that responds to proteins in human cells.....	III-1
Abstract.....	III-2
Introduction.....	III-3
Results.....	III-7
Design of protein-responsive ribozyme switch platforms.....	III-7
Initial characterization of protein-responsive ribozyme switches in a human cell line	III-10
Development of an improved genetic system for quantitative characterization of ribozyme switch activity <i>in vivo</i>	III-15
MS2 variants result in optimization of switch sensitivity to ligand.....	III-21
Protein ligand localization allows probing of ribozyme switch mechanism of action	III-24
Development of ribozyme switch platforms responsive to additional proteins....	III-34
Discussion.....	III-43
Methods.....	III-49
Acknowledgements.....	III-60
References.....	III-62
Supplementary Tables.....	III-67
CHAPTER 4: Conclusions.....	IV-1
References.....	IV-6

List of Figures

- Figure 1.1 The phosphodiester isomerization mechanism of hammerhead ribozymes
- Figure 1.2 Assembly of a ribozyme switch from modular components
- Figure 1.3 Ligand binding stabilizes the aptamer-formed conformation
-
- Figure 2.1 Structure and function of the trans-ribozyme
- Figure 2.2 Structures of trans-ribozymes bound to target sequences
- Figure 2.3 Trans-ribozyme and target gene characterization system
- Figure 2.4 Activity of trans-ribozymes
- Figure 2.5 Structures of cis-ribozymes used to model trans-ribozyme activity
- Figure 2.6 Activity of type I cis-ribozymes
- Figure 2.7 The improved trans-ribozyme
- Figure 2.8 Trans-ribozyme ancillary elements
- Figure 2.9 Activity of trans-ribozymes with ancillary elements
- Figure 2.10 Plasmid maps
-
- Figure 3.1 MS2-responsive ribozyme switch designs
- Figure 3.2 Protein-responsive ribozyme switch characterization system
- Figure 3.3 Activity of MS2-A designs
- Figure 3.4 Activity of MS2-B, MS2-C, and MS2-D designs
- Figure 3.5 Histograms of transiently transfected and stably integrated fluorescent constructs
- Figure 3.6 Effects of transient transfection of stable cell lines expressing a fluorescent reporter protein

- Figure 3.7 Activity of ribozyme switches with optimized ligands
- Figure 3.8 Subcellular localization of ligands with and without localization signals
- Figure 3.9 Activity of ribozyme switches with ligands with and without localization signals
- Figure 3.10 Confocal fluorescence microscopy of ligands with and without localization signals
- Figure 3.11 Activity of ribozyme switches with improved localized ligands
- Figure 3.12 Activity of ribozyme switches with improved localized ligands over a range of ligand concentrations
- Figure 3.13 Design and testing of PP7-responsive switch designs
- Figure 3.14 Design and testing of lambda-N-responsive switch designs
- Figure 3.15 NF- κ B-responsive ribozyme switch characterization system
- Figure 3.16 Design and testing of NF- κ B-responsive switch designs
- Figure 3.17 Design and testing of β -catenin-responsive switch designs
- Figure 3.18 Plasmid maps

Chapter 1

Introduction

RNA is a versatile regulatory biomolecule

Synthetic biology is a rapidly emerging field that promises to improve our ability to investigate and manipulate living organisms through the creation of novel biological tools and systems, with innovations supporting applications in health, energy, and biomanufacturing¹⁻³. While advances in DNA synthesis have enabled the construction of large genetic systems⁴, the capability to design and predictably regulate such systems lags behind. Synthetic RNA-based gene-regulatory devices are uniquely poised to address this need.

Once thought to be merely the intermediate between the genetic information stored in DNA and proteins that executed cellular function, RNA has been shown to perform a large diversity of functional activities, such as catalysis, metabolite binding, and gene regulation⁵⁻⁸. In addition, functional RNA molecules have been described that can modulate their activity in response to cellular and environmental inputs. For example, temperature-sensitive structural elements regulate gene expression in the heat and cold shock responses in bacteria⁹, and metabolite-binding elements control the expression of enzymes in biosynthetic pathways¹⁰⁻¹². To date most of these regulatory elements have been characterized in prokaryotes, but examples have been found in eukaryotes as well¹³. The many examples of naturally-occurring, ligand-responsive RNA-based gene-regulatory elements, or RNA switches, serve as the raw materials and inspiration for novel synthetic RNA-based regulatory devices¹⁴.

As with proteins, the ability of RNA to perform functional activities arises from its three-dimensional folded structure. Unlike proteins, however, this structure is almost entirely determined by hydrogen-bonding, base-stacking, and electrostatic interactions

between the constituent monomers¹⁵. The relative simplicity of RNA intramolecular interactions has enabled the design of software models that computationally predict the secondary structures and associated free energies of a given RNA sequence with a high degree of accuracy¹⁶⁻¹⁸. Such software has greatly aided the design of engineered functional RNA molecules^{19,20}. Facile protein structure prediction is not yet feasible due to the complexity of protein folding, and therefore protein-based devices such as allosteric transcription factors are currently far more challenging to engineer than their RNA-based counterparts.

Engineered RNA devices in eukaryotes enable dynamic modulation of gene expression in response to molecular and environmental signals

Synthetic RNA switches achieve gene regulation through a variety of mechanisms, but they generally contain two core components. The sensor component detects the input signal, such as a small molecule or protein, through a binding interaction, and the actuator component modulates gene expression through mechanisms such as transcription, post-transcriptional processing, translation, or messenger RNA (mRNA) stability. Many RNA-based devices utilize architectures that also incorporate a transmitter component, which links the sensor and actuator components and transmits information between them by modulating the activity of the actuator based on the ligand bound state of the sensor. The sensor component is typically an aptamer, an RNA sequence with high affinity and specificity for a small molecule or protein ligand. Many such binding elements can be found in nature^{12,21}, but new aptamers can be generated

with an *in vitro* selection method known as systematic evolution of ligands by exponential enrichment, or SELEX^{22,23}. This method can be used to generate aptamer sequences to theoretically any small molecule or protein ligand of interest.

The earliest potential point of regulation of gene expression is transcription. In one example, an RNA regulator of transcription that responded to the small molecule tetramethylrosamine (TMR) was demonstrated in *Saccharomyces cerevisiae*. The TMR aptamer was linked to a transcriptional activator through a randomized transmitter component and functional devices were selected based on TMR responsiveness. Demonstrations of engineered ligand-responsive RNA-based regulators of transcription have not been reported to date in mammalian cells.

RNA-based devices that modulate gene expression through post-transcriptional processing, such as splicing, have been demonstrated in yeast and human cells^{24,25}. Proper assembly of the spliceosome requires recognition of specific exonic and intronic sequence elements, and researchers have shown that the accessibility of these elements can be regulated by ligand binding to aptamer sequences. In one example, an aptamer for tetracycline was placed at the 5' splice site in the precursor mRNA (pre-mRNA) of a fluorescent reporter gene in yeast²⁰. Binding of tetracycline altered the conformation of the region around the splice site, preventing splicing of the exons encoding the reporter gene and reducing expression by up to 32-fold. In another example, protein-responsive RNA-based devices were used to control alternative splicing of different transgenes in human embryonic kidney 293 (HEK293) cells²⁵. Aptamers for three different proteins were placed in an intronic region such that protein binding to the aptamer sequences

prevented the exclusion of an exon containing a premature stop codon, thereby modulating the expression of the encoded transgene.

RNA interference (RNAi) is another post-transcriptional processing mechanism that has been utilized in RNA-based devices for controlling target gene expression. RNAi is a powerful platform for gene regulation in higher eukaryotes that is based on complementarity between the RNA regulator and the target gene, where the regulators can be encoded in diverse forms including microRNAs (miRNAs), short hairpin RNAs (shRNAs), and small interfering RNAs (siRNAs)²⁶. These RNAi-based components must be processed by cellular protein machinery to silence gene expression, either through blocking translation initiation, interrupting polypeptide elongation, or degrading the transcript¹⁴. Many RNA switches that modulate processing in response to ligand input have been demonstrated^{27,28}. In one example, a miRNA-based switch responsive to small molecules was demonstrated in HEK293 cells²⁹. Ligand binding to an aptamer integrated into the base of the miRNA stem prevented processing of the primary miRNA (pri-miRNA) by Drosha, thereby increasing target gene expression levels as a function of increasing ligand concentrations. In another example, the aptamer for the archaeal ribosomal protein L7Ae was inserted in the loop region of an shRNA targeting an antiapoptotic gene²⁷. By simultaneously regulating a proapoptotic gene with a separate device, the authors were able to control apoptosis in HeLa cells.

Regulation of translation initiation is a common mechanism employed by ligand-responsive RNA switches. Following the example of natural prokaryotic translation initiation riboswitches and their engineered counterparts¹⁰⁻¹², the aptamer is placed in the 5' untranslated region (UTR) just upstream of the translation initiation codon, such that

ligand binding prevents the ribosome from binding and assembling properly. For switches responsive to small molecules³⁰, ligand binding can stabilize structures that discourage ribosome assembly, while in other cases protein binding prevents ribosome association through steric hindrance³¹⁻³³. In one interesting study in HEK293T cells, protein binding to an aptamer in the 5' UTR of a bicistronic mRNA selectively repressed translation of the upstream gene while not affecting internal ribosome entry sequence (IRES)-dependent translation of the downstream gene³².

Finally, effective regulation of gene expression can be accomplished by controlling the stability of mRNA, usually by modulating the susceptibility of mRNA to cellular ribonucleases (RNases). The ends of eukaryotic mRNAs are protected by the 5' 7-methyl-guanosine cap and the 3' poly(A) tail, which themselves are bound by various proteins that circularize the transcript. Directed cleavage in either of the UTRs or the coding region exposes the mRNA to rapid degradation by exoribonucleases. In one engineered switch exploiting this phenomenon, an aptamer that binds the caffeine analogue theophylline was integrated into a hairpin recognized by the RNase Rnt1p, such that ligand binding prevented Rnt1p-mediated cleavage in yeast³⁴. Another type of device controlling mRNA stability is based on self-cleaving ribozymes, which will be described below.

Ligand-responsive ribozyme switches

Ribozymes are RNA enzymes that accelerate chemical reactions by adopting certain folded structures similar to peptide-based enzymes. Natural ribozymes were first

discovered in Group I introns³⁵, but have since been identified to be involved in many vital cellular processes from mRNA splicing³⁶ to peptide synthesis³⁷. Many ribozymes catalyze the lysis of an RNA phosphodiester bond, either in its own strand (cis) or in a separate RNA molecule (trans), thereby cleaving it in two. Hammerhead ribozymes, first discovered in plant viroids³⁸ and shown to function in a variety of organisms³⁹, rapidly catalyze self-cleavage through a phosphodiester isomerization mechanism (Figure 1.1). The cleavage site is located in the ribozyme's catalytic core immediately downstream of the conserved NUX sequence, in which N is any nucleotide and X is either A, C, or U.

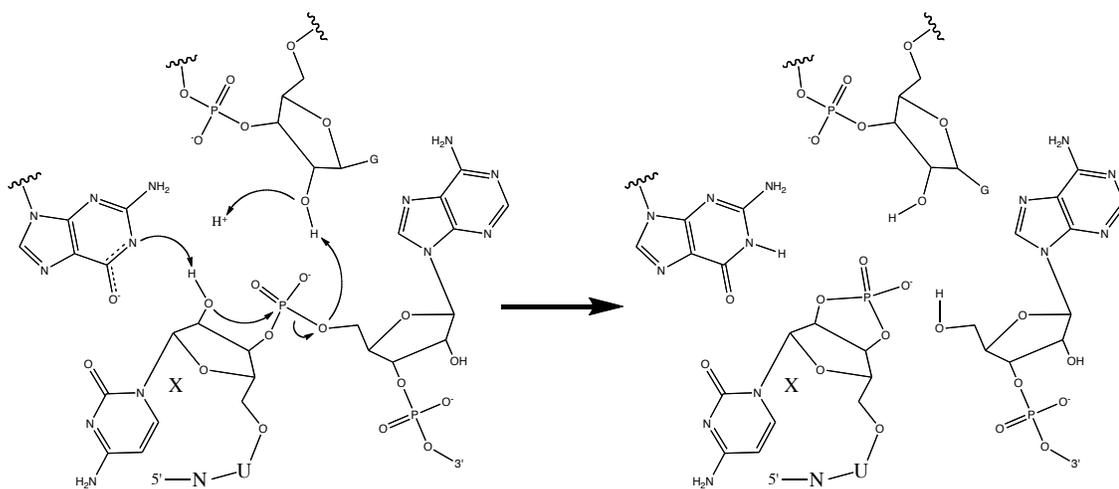


Figure 1.1. The phosphodiester isomerization mechanism of hammerhead ribozymes. Two nearby guanosines contribute to general base catalysis. In this example, 'X' is cytosine. Adapted from⁴⁰.

The Smolke laboratory has recently described a framework for constructing ribozyme-based gene-regulatory RNA devices¹⁹. The framework provides a modular assembly strategy for building these RNA devices from a sensor component, made of an

RNA aptamer, an actuator component, made of a satellite RNA of tobacco ringspot virus (sTRSV) ribozyme⁴¹, and a transmitter component, made of a sequence that functionally couples the sensor and actuator components (Figure 1.2). The transmitter component is rationally designed based on competitive hybridization events that enable the device to distribute between two primary conformations: one in which the input cannot bind to the sensor and the other in which the input can bind to the sensor. Input binding shifts the distribution to favor the input-bound conformation as a function of increasing input concentration and is translated to a change in the activity of the actuator, where a ‘ribozyme-active’ state results in self-cleavage of the device. The RNA device is coupled to the 3’ UTR of the target gene, where ribozyme self-cleavage inactivates the transcript and thereby lowers gene expression independent of cell-specific machinery.

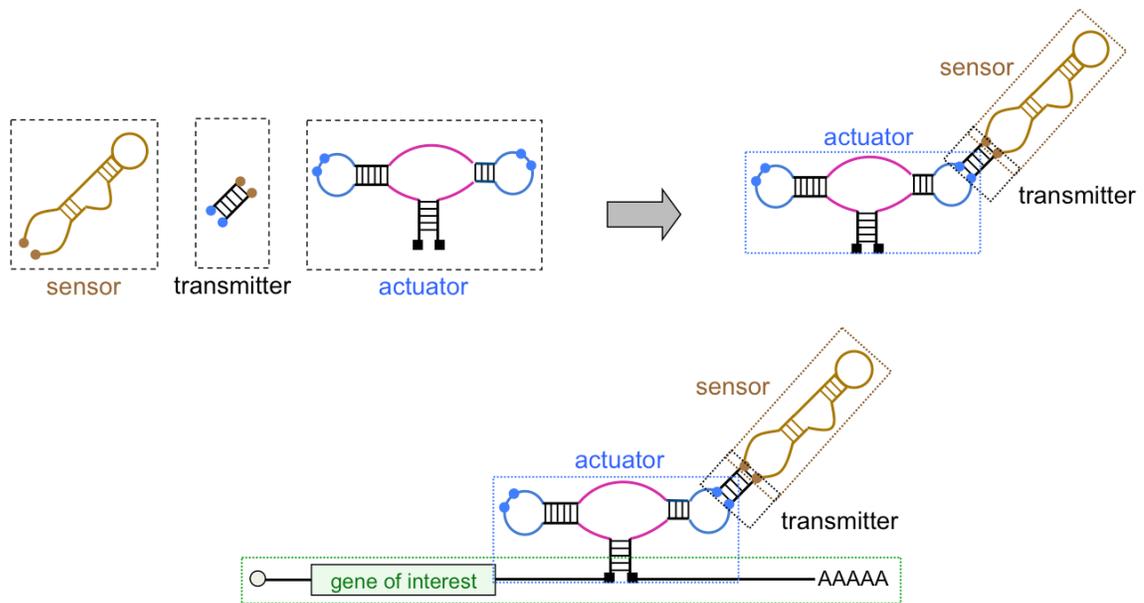


Figure 1.2. Assembly of a ribozyme switch from modular components. The aptamer is shown in light brown, the stems are shown in black, the catalytic core is shown in magenta, and loops and bulges are shown in blue. Adapted from⁴².

The precise design of the transmitter component determines whether the ribozyme switch will repress or enhance gene expression, unlike many of the switches described above, which are capable of regulating gene expression in only one direction. RNA devices that function as either ON or OFF switches that convert a molecular input signal to increased or decreased gene expression output, respectively, have been demonstrated in yeast and mammalian cells^{19,42-46} (Figure 1.3). After initial demonstration of ribozyme switches responsive to theophylline and tetracycline in yeast¹⁹, the framework was extended to provide a general approach for the engineering of multi-input, higher-order information processing devices, where two-input logic gates (AND, NOR, NAND, and OR gates), signal filters, band-pass filters, and programmed cooperativity operations were demonstrated⁴². These ribozyme switches were also used to control T-cell proliferation in mice⁴³, demonstrating phenotypic control in an animal model. Other investigators have demonstrated switching activity of a theophylline-responsive ribozyme switch coupled to the 5' UTR^{45,46}, but this strategy can lead to nonspecific reduction of translation initiation due to the high degree of secondary structure upstream of the start codon.

Ribozyme switches possess a significant advantage not shared by many other gene regulation platforms in that their mechanism of action does not require any cell-specific machinery. Ribozyme switches are therefore functional across different organisms, including bacteria⁴⁷, yeast¹⁹, and mammalian systems⁴³. This allows rapid screening of devices generated by both rational and directed evolution design strategies in simple organisms⁴⁸, optimizing device activity before transitioning to more complex organisms⁴⁴.

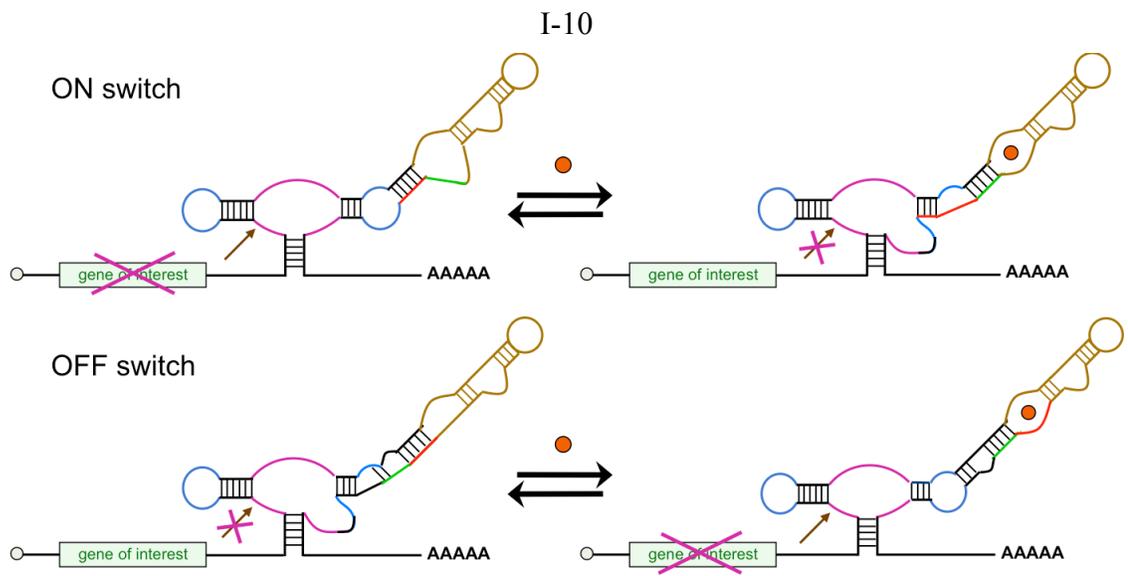


Figure 1.3. Ligand binding stabilizes the aptamer-formed conformation. In an ON switch, ligand (red disk) binding stabilizes the catalytically inactive conformation, preventing ribozyme self-cleavage and allowing translation of the gene of interest. In an OFF switch, the ligand stabilizes the catalytically active conformation, inducing cleavage and gene repression. The cleavage site is indicated with an arrow. Coloring is the same as in Figure 1.2.

However, ribozyme switches are somewhat limited in their effectiveness and range of capabilities. First, they are generally limited to the regulation of transgenes, with endogenous gene regulation achievable only through the utilization of targeted chromosomal integration strategies, which are cumbersome in mammalian systems^{49,50}. In contrast, switch platforms based on RNAi enable facile ligand-responsive regulation of endogenous genes^{27,28}. Second, to date only small-molecule-responsive ribozyme switches have been described, while other platforms have been shown to respond to protein ligands. Third, ribozyme switches have not yet been able to achieve the high dynamic ranges and input sensitivities of other gene regulation systems. Finally, the

mechanism of action, specifically the subcellular location where ribozyme cleavage occurs, has not been fully elucidated. For other switch platforms, such as those based on modulation of transcription, splicing, or RNAi processing, choice of ligand is constrained by the known subcellular location of the mechanism of action. It is desirable to elucidate similar details about ribozyme switches to determine which ligands the platform is capable of sensing.

Applications

Engineered RNA devices have been used for a variety of applications in eukaryotes. In reconstituting useful biosynthetic pathways in new host organisms, it is important to regulate the expression levels of the enzymes to maximize their activity while efficiently exploiting cellular resources. Ribozyme-based regulatory devices have been used as noninvasive sensors of enzymatic products. In one example in yeast, a ribozyme switch responsive to xanthine was used to control a fluorescent reporter gene¹⁹. When the yeast were fed xanthosine, the enzymatic conversion of xanthosine to xanthine was reported noninvasively by fluorescent output. In an extension of this concept, a theophylline-responsive ribozyme switch controlling a fluorescent reporter gene was used in a high-throughput screen of a large enzyme library of a caffeine demethylase, identifying a variant with 33-fold improvement in catalytic activity over eight rounds of directed evolution⁵¹.

Synthetic RNA switches have demonstrated applications for medical purposes in human cells. In one notable example, an RNA switch controlling alternative splicing

modulated protein expression levels in response to nuclear factor κ B (NF- κ B) and β -catenin, two signaling proteins with important roles in disease²⁵. The device was able to influence cell fate by controlling the levels of a gene conferring sensitivity to a drug that induces apoptosis. In another example, ribozyme switches responsive to small molecule drugs were used to regulate the expression of the cytokines IL-2 and IL-15 in engineered T cells, thereby imparting drug-modulated control over T-cell proliferation and survival *in vitro* and *in vivo*⁴³. This latter system was demonstrated in a mouse model, highlighting the potential application of this technology to improving the safety and efficacy of adoptive immunotherapy strategies.

The future holds many more potential applications in biosensors, biofuels and drug compounds from synthetic metabolic pathways, diagnostic tools, and next-generation gene therapies. Additionally, all new applications, as well as all of the demonstrations described above, provide insight into the underlying biological mechanisms on which they rely, increasing our understanding of natural systems and how to better manipulate those systems in the future.

Scope of thesis

This thesis describes the development of a synthetic RNA device platform for the regulation of gene expression in response to molecular signals. As described in Chapter 2, we began with the cis-acting ribozyme switch platform developed by Win and Smolke¹⁹, attempting to divide the structure into two RNA strands such that their annealing would reconstitute the functional device. These trans-ribozyme-based devices

were designed to target synthetic sequences inserted into the 3' UTR of the target gene and were expressed in human cells. After optimizing the molecular design for maximal *in vivo* cleavage activity using a cis-ribozyme-based model system, the improved trans-ribozyme was coupled with additional RNA elements intended to increase the likelihood of binding and cleavage of the target strand. However, *in vivo* activity of trans-ribozymes was not established, likely due to the inability of the two RNA strands to properly hybridize inside the cell. Chapter 3 describes the development of protein-responsive ribozyme switches. We designed a variety of device architectures intended to respond to the bacteriophage MS2 coat protein through different switching mechanisms. We developed a genetic system for quantitative characterization of the activity of these devices in human cells. After demonstrating a range of regulatory capabilities among the various device designs, we investigated the impact of different MS2 subcellular localizations on device activity and found that the switch platform is able to respond to both cytoplasmic- and nuclear-localized ligand. Finally, we designed ribozyme switches to respond to other protein ligands in order to demonstrate the versatility of our device platform. Chapter 4 discusses future directions for this work and its contributions to the field.

References

1. Khalil, A. & Collins, J. Synthetic biology: applications come of age. *Nat. Rev. Genet.* **11**, 367–379 (2010).
2. Chen, Y. Y. et al. Synthetic biology: advancing biological frontiers by building synthetic systems. *Genome Biol.* **13**, 240 (2012).
3. Chen, Y. Y. & Smolke, C. D. From DNA to targeted therapeutics: bringing synthetic biology to the clinic. *Sci. Transl. Med.* **3**, 106ps42 (2011).
4. Ma, S. et al. DNA synthesis, assembly and applications in synthetic biology. *Curr. Opin. Chem. Biol.* **16**, 260–267 (2012).
5. Chang, A. L. et al. Synthetic RNA switches as a tool for temporal and spatial control over gene expression. *Curr. Opin. Biotechnol.* **23**, 679–688 (2012).
6. Thodey, K. & Smolke, C. D. Bringing it together with RNA. *Science.* **333**, 412–413 (2011).
7. Liang, J. C. et al. Engineering Biological Systems with Synthetic RNA Molecules. *Mol. Cell* **43**, 915–926 (2011).
8. Isaacs, F. J. et al. RNA synthetic biology. *Nat. Biotechnol.* **24**, 545–554 (2006).
9. Kozak, M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* **361**, 13–37 (2005).
10. Winkler, W. C. & Breaker, R. R. Genetic control by metabolite-binding riboswitches. *Chembiochem* **4**, 1024–1032 (2003).
11. Mandal, M. & Breaker, R. R. Gene regulation by riboswitches. *Nat. Rev. Mol. Cell Biol.* **5**, 451–463 (2004).
12. Winkler, W. C. & Breaker, R. R. Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.* **59**, 487–517 (2005).
13. Sudarsan, N., Barrick, J. & Breaker, R. Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA* **9**, 644–647 (2003).
14. Stefani, G. & Slack, F. J. Small non-coding RNAs in animal development. *Nat. Rev. Mol. Cell Biol.* **9**, 219–230 (2008).
15. Win, M. N. et al. Frameworks for programming biological function through RNA parts and devices. *Chem. Biol.* **16**, 298–310 (2009).

16. Mathews, D. H. Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *RNA* **10**, 1178–1190 (2004).
17. Mathews, D. H. & Turner, D. H. Prediction of RNA secondary structure by free energy minimization. *Curr. Opin. Struct. Biol.* **16**, 270–278 (2006).
18. Zadeh, J. et al. NUPACK: analysis and design of nucleic acid systems. *J. Comput. Chem.* **32**, 170–173 (2011).
19. Win, M. N. & Smolke, C. D. A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14283–14288 (2007).
20. Kennedy, A. B. et al. A versatile cis-blocking and trans-activation strategy for ribozyme characterization. *Nucleic Acids Res.* **41**, e41 (2013).
21. Tucker, B. J. & Breaker, R. R. Riboswitches as versatile gene control elements. *Curr. Opin. Struct. Biol.* **15**, 342–348 (2005).
22. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*. **249**, 505–510 (1990).
23. Ellington, A. & Szostak, J. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990).
24. Weigand, J. E. & Suess, B. Tetracycline aptamer-controlled regulation of pre-mRNA splicing in yeast. *Nucleic Acids Res.* **35**, 4179–4185 (2007).
25. Culler, S. J. et al. Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins. *Science*. **330**, 1251–1255 (2010).
26. Ramachandran, P. V. & Ignacimuthu, S. RNA interference—a silent but an efficient therapeutic tool. *Appl. Biochem. Biotechnol.* **169**, 1774–1789 (2013).
27. Saito, H. et al. Synthetic human cell fate regulation by protein-driven RNA switches. *Nat. Commun.* **2**, 160 (2011).
28. Beisel, C. L. et al. Model-guided design of ligand-regulated RNAi for programmable control of gene expression. *Mol. Syst. Biol.* **4**, (2008).
29. Beisel, C. L. et al. Design of small molecule-responsive microRNAs based on structural requirements for Drosha processing. *Nucleic Acids Res.* **39**, 2981–2994 (2011).

30. Kötter, P. et al. A fast and efficient translational control system for conditional expression of yeast genes. *Nucleic Acids Res.* **37**, e120 (2009).
31. Paraskeva, E. et al. A translational repression assay procedure (TRAP) for RNA-protein interactions in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 951–956 (1998).
32. Nie, M. & Htun, H. Different modes and potencies of translational repression by sequence-specific RNA-protein interaction at the 5'-UTR. *Nucleic Acids Res.* **34**, 5528–5540 (2006).
33. Saito, H. et al. Synthetic translational regulation by an L7Ae-kink-turn RNP switch. *Nat. Chem. Biol.* **6**, 71–78 (2010).
34. Babiskin, A. H. & Smolke, C. D. Engineering ligand-responsive RNA controllers in yeast through the assembly of RNase III tuning modules. *Nucleic Acids Res.* **39**, 5299–5311 (2011).
35. Grabowski, P. J. et al. The intervening sequence of the ribosomal RNA precursor is converted to a circular RNA in isolated nuclei of *Tetrahymena*. *Cell* **23**, 467–476 (1981).
36. Jacquier, A. & Rosbash, M. Efficient Trans-Splicing of a Yeast Mitochondrial RNA Group II Intron Implicates a Strong 5' Exon-Intron Interaction. *Science.* **234**, 1099–1104 (1986).
37. Rodnina, M. V. et al. How ribosomes make peptide bonds. *Trends Biochem. Sci.* **32**, 20–26 (2007).
38. Prody, G. et al. Autolytic Processing of Dimeric Plant Virus Satellite RNA. *Science.* **231**, 1577–1580 (1986).
39. Birikh, K. R. et al. The structure, function and application of the hammerhead ribozyme. *Eur. J. Biochem.* **245**, 1–16 (1997).
40. Chi, Y.-I. et al. Capturing hammerhead ribozyme structures in action by modulating general base catalysis. *PLoS Biol.* **6**, e234 (2008).
41. Flores, R. et al. Hammerhead ribozyme structure and function in plant RNA replication. *Methods Enzymol.* **341**, 540–552 (2001).
42. Win, M. N. & Smolke, C. D. Higher-order cellular information processing with synthetic RNA devices. *Science.* **322**, 456–460 (2008).
43. Chen, Y. Y. et al. Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 8531–8536 (2010).

44. Wei, K. Y. et al. A yeast-based rapid prototype platform for gene control elements in mammalian cells. *Biotechnol. Bioeng.* **110**, 1201–1210 (2013).
45. Ausländer, S. et al. A ligand-dependent hammerhead ribozyme switch for controlling mammalian gene expression. *Mol. Biosyst.* **6**, 807–814 (2010).
46. Wieland, M. et al. Engineering of ribozyme-based riboswitches for mammalian cells. *Methods* **56**, 351–357 (2012).
47. Wieland, M. & Hartig, J. S. Improved aptazyme design and in vivo screening enable riboswitching in bacteria. *Angew. Chemie* **47**, 2604–2607 (2008).
48. Liang, J. C. et al. A high-throughput, quantitative cell-based screen for efficient tailoring of RNA device activity. *Nucleic Acids Res.* **40**, e154 (2012).
49. Urnov, F. D. et al. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636–646 (2010).
50. Bogdanove, A. J. & Voytas, D. F. TAL effectors: customizable proteins for DNA targeting. *Science.* **333**, 1843–1846 (2011).
51. Michener, J. K. & Smolke, C. D. High-throughput enzyme evolution in *Saccharomyces cerevisiae* using a synthetic RNA switch. *Metab. Eng.* **14**, 306–316 (2012).

Chapter 2

Development of an RNA device framework that targets endogenous genes in human cells

Abstract

Ligand-responsive genetic control systems are important tools in synthetic biology. Such tools are especially valuable when they include the capability to regulate endogenous genes. Allosteric ribozyme switches have been designed based on hammerhead ribozymes and RNA aptamers, and have demonstrated programmable ligand-responsive genetic regulation in diverse cell types. We attempted to adapt this class of cis-acting genetic control elements to function in trans. Previous work has demonstrated the division of a cis-acting hammerhead ribozyme into an enzyme strand and a substrate strand that reconstitute catalytic activity upon annealing with one another. We developed a design strategy to divide the allosteric ribozyme switch into two strands, such that the sensor component is entirely contained within the enzyme strand. We investigated the ability of our trans-ribozyme designs to regulate the expression of genes in trans in human cell lines. Cleavage activity of the trans-ribozyme platform was optimized using cis-ribozymes as a model, and our results indicate that the ribozyme stem sequence is not as mutable as previously reported. We verified the cleavage activity of our optimized trans-ribozyme design *in vitro*, and coupled that design to a variety of ancillary genetic elements to direct stability, structure, processing, and localization of the ribozyme transcript *in vivo*. However, we were unable to demonstrate trans-ribozyme-mediated gene silencing, likely due to deficiencies in trans-ribozyme transcript stability and localization.

Introduction

The ability to regulate the expression of endogenous genes is a desired function for synthetic RNA-based control systems. The capability to interact with and modulate endogenous genes enables the silencing of the negative effects of gene products from pathogenic RNA and aberrant messenger RNA (mRNA), forming the foundation for novel gene therapies and tissue engineering methodologies. Such targeted gene silencing has been demonstrated in models of bacterial infection¹, viral infection²⁻⁷, and cancer⁸⁻¹¹. For example, ribozymes have been used to target multiple genes in the HIV genome, effectively inhibiting viral replication in both laboratory studies^{12,13} and clinical trials¹⁴⁻¹⁶. In another example, tumor growth and angiogenesis in a pancreatic cancer mouse model were inhibited by a short hairpin RNA (shRNA) targeting glycogen synthase kinase-3 β , an important serine/threonine protein kinase in tumorigenesis¹¹.

When regulating genes in mammalian cells using synthetic RNA devices, it is often desirable to control the activity of those devices in response to user-specified molecular inputs. This is especially true in the case of cancer therapeutics, where an important strategy to increase the efficacy and safety of the therapy is to target the regulatory effect to diseased cells while leaving healthy cells unaffected. Such ligand-responsive RNA-based genetic control elements have been demonstrated in mammalian cells. In one example, alternative splicing was modulated using switches responsive to cancer biomarkers, such that presence of the biomarker allowed expression of herpes simplex virus-thymidine kinase (HSV-TK), conferring sensitivity to the pro-drug ganciclovir¹⁷. In another example, the balance between proapoptotic and antiapoptotic

genes was controlled using shRNAs containing an aptamer for the archaeal ribosomal protein L7Ae, whose processing was inhibited by ligand binding¹⁸.

There are several desirable features for an effective ligand-responsive gene-regulatory device. Many previously described platforms exhibit some of these key features, but very few exhibit all of them. First, the device must be programmable to respond to different ligand inputs, turning gene expression either on or off in response to ligand binding. Many of the reported ligand-responsive platforms are capable of modulating gene expression either up or down, but not both¹⁹⁻²². Second, the basal level of activity and the switching range of the device must be readily tunable through small alterations to the design to easily adjust device function to application-specific levels. Third, the ligand sensor and gene-regulatory actuator components must be modular in assembly, such that the ligand-binding domain can be easily replaced with a sensor for a different input, and the actuator can be retargeted to regulate a different gene, without necessitating a full and lengthy redesign of the device. Lastly, a device platform that is portable between organisms, such as microbes and higher eukaryotes, can allow for rapid prototyping and optimization of the device in simple organisms and later implementation in more complex organisms. This property is limited to devices that incorporate actuators that do not depend on cell-specific machinery.

RNA control elements derived from the hammerhead ribozyme of the satellite RNA of tobacco ringspot virus (sTRSV)²³ have been demonstrated to exhibit these desired capabilities and thus provide a powerful ligand-responsive platform for mammalian gene regulation. The allosteric ribozyme switch framework developed by Win and Smolke²⁴ demonstrates programmable ligand-responsive genetic regulation

through a synthetic RNA device. These devices transmit ligand sensing by an aptamer component into cleavage of the target gene's mRNA by a ribozyme actuator component, which leads to degradation of the transcript and silencing of gene expression²⁴. The ribozyme switches can be programmed to respond to different ligand inputs through the incorporation of different aptamer sequences²⁴. The activity of ribozyme switches is readily tuned by altering individual nucleotides, which changes the three-dimensional folded state of the device, thus altering the basal level of catalytic activity and the energy difference between the active and inactive conformations. This, in turn, determines the difference in gene expression between the ON and OFF states^{24,25}. The modular components of the ribozyme switch platform can be easily replaced without affecting device activity, and the switch can be placed in the 3' untranslated region (UTR) of any gene of interest to regulate its expression²⁴. Finally, because ribozyme cleavage does not rely on any cell-specific machinery, the platform is highly portable between organisms, supporting rapid prototyping systems that allow designs to be screened in a microbial host such as yeast and optimized designs subsequently ported to mammalian cells with little change in function²⁶.

The primary limitation of the ribozyme switch platform, as with many other previously demonstrated ligand-responsive regulation devices, is that it cannot be used to control endogenous genes^{24,27}. Instead it is limited to the regulation of transgenes, as the cis-acting genetic actuator must be encoded in the region immediately neighboring the target gene. However, previous work has demonstrated that the hammerhead ribozyme can function as two separate molecules, an enzyme strand and a substrate strand, that reconstitute catalytic activity upon annealing with one another^{28,29}. These trans-acting

ribozymes can be tailored to a specific target sequence through the identity of targeting arms that base-pair to regions in the target gene, and can thus be used to target endogenous genes. In an early demonstration, a trans-ribozyme was programmed to target the *gag* gene of HIV-1, lowering levels of that transcript in human cells².

Since this initial demonstration of endogenous gene regulation, investigators have examined factors that determine the functional activity of trans-ribozymes *in vivo*. Trans-ribozymes were found to function far more effectively in the cytoplasm than the nucleus, and localization strategies have been employed to target trans-ribozyme transcripts to the cytoplasm^{30,31}. Taira and colleagues coupled trans-ribozymes to transfer RNA (tRNA) to take advantage of its cytoplasmic localization and stability^{29,31}, and used a random library to screen the region linking the trans-ribozyme and tRNA for increased stability³². Another important factor is the secondary structure of both the trans-ribozyme and target transcripts, which can interfere with binding. In one notable study, a trans-ribozyme was linked to an RNA helicase protein, which removed secondary structure from the target mRNA to allow proper binding and cleavage³³. However, there is disagreement in the field on the effectiveness of trans-ribozymes as gene-regulatory elements, as the studies on trans-ribozymes have rarely included a non-cleaving control trans-ribozyme to clearly demonstrate that observed levels of gene expression knockdown are due to mRNA cleavage from the ribozyme, rather than antisense effects as a result of binding of the trans-ribozyme to the transcript. Indeed, one study investigating trans-ribozymes found that these gene-regulatory elements were no more effective at silencing their target gene than equivalent non-catalytic antisense sequences³.

We attempted to extend the cis-acting ribozyme switch platform developed by Win and Smolke²⁴ to a trans-acting platform capable of regulating the expression of endogenous genes. We divided the allosteric ribozyme switch into an enzyme strand and a substrate strand, such that the sensor component is entirely contained within the enzyme strand. We sought to leverage all of the advantages of the existing cis-ribozyme switch platform while overcoming its limitation of being able to regulate only heterologous genes. We designed three trans-ribozymes and placed their cognate target sequences in the 3' UTR of a fluorescent reporter gene, which we integrated into the chromosome of a human cell line to model the targeting of an endogenous gene. Based on our initial results indicating that the trans-ribozyme designs were unable to silence the target gene, we performed additional studies to optimize the cleavage activity and gene expression knockdown in a model cis-ribozyme architecture, which led to the development of an improved trans-ribozyme design. We also varied the trans-ribozyme expression system, incorporating genetic elements intended to increase the ability of the trans-ribozyme to anneal to and cleave the target strand. Our results indicate that the sequence flexibility of the trans-ribozyme is severely restricted, limiting the capability to design trans-ribozymes to target any gene of choice. We were unable to demonstrate gene regulation *in vivo* from our trans-ribozyme designs, likely due to issues of trans-ribozyme transcript stability and localization.

Results

Design of a trans-ribozyme-based regulatory element in human cells

We first attempted to establish the capabilities of trans-acting RNA devices to regulate endogenous cellular transcripts in human cells. The design of the trans-ribozyme is based on a modification of a hammerhead ribozyme³⁴ that was optimized to enhance cleavage activity in the presence of physiological Mg^{2+} concentrations and hybridization efficiency between the two strands. As shown in Figure 2.1A, the sequence of the hammerhead ribozyme is divided in two at loop I, such that the cleavage site is located in the target transcript. Stems I and III are formed through the hybridization of the ribozyme targeting arms to complementary regions of the target transcript, whereas stem II and loop II are entirely contained within the trans-acting ribozyme strand. This places almost all of the nucleotides reported to be conserved in the enzyme strand, with only the conserved NUX cleavage site in the target strand³⁵. Such designs have shown higher cleavage activity *in vitro* than designs in which the ribozyme is divided at loop II³⁴, and they are more directly adapted to the cis-ribozyme-based RNA device framework²⁴, as any aptamer can then be integrated into loop II. To maintain the tertiary interactions between nucleotides in loops I and II that have been shown to be necessary for catalytic activity at physiological Mg^{2+} concentrations³⁶, stem I of the ribozyme strand contains a bulge that mimics loop I. When the ribozyme strand and target transcript anneal the catalytic core is effectively reconstituted and the target strand is cleaved. Integration of the target sequence in the flexible regulatory space of the 3' UTR of a reporter gene enables knockdown of that gene through targeted cleavage and subsequent degradation of its mRNA (Figure 2.1B).

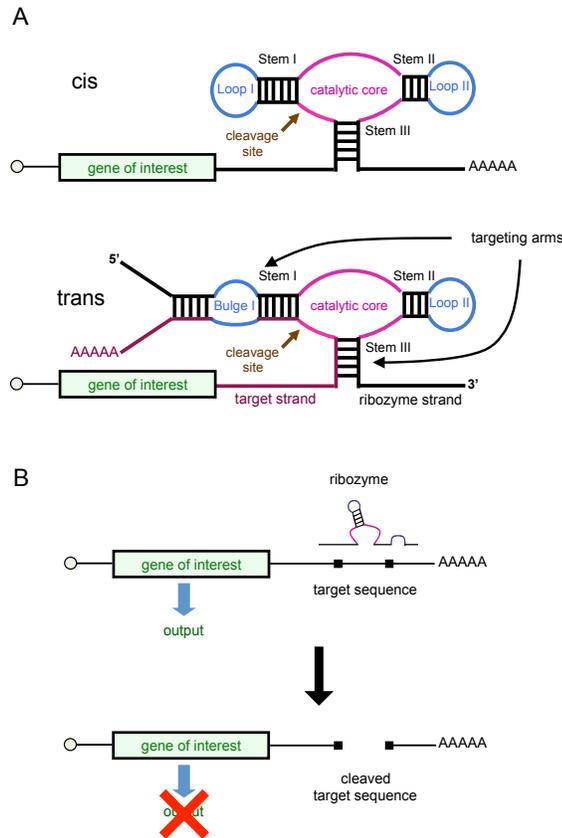


Figure 2.1. Structure and function of the trans-ribozyme. **(A)** The hammerhead ribozyme in cis and trans forms. The catalytic core is shown in magenta, loops and bulges are shown in blue, the ribozyme strand is shown in black, and the target strand is shown in purple. The cleavage site is indicated with an arrow. **(B)** The trans-ribozyme binds and cleaves the target sequence in the 3' UTR of the gene of interest, destabilizing the transcript and reducing protein expression. Partially adapted from Win and Smolke²⁴.

Preliminary studies previously performed in the Smolke laboratory focused on the optimization of trans-ribozyme activity under physiological conditions. *In vitro* experiments on a trans-ribozyme derived from the sTRSV hammerhead ribozyme demonstrated that the length of the targeting arms significantly impacts hybridization

interactions, and therefore cleavage rate, at physiological Mg^{2+} concentrations. Specifically, cleavage activity was shown to be highest when the targeting arm 5' of bulge I and the stem III targeting arm are 16 and 7 base pairs long, respectively. Preliminary experiments conducted in yeast showed limited trans-ribozyme activity (Kate Galloway, unpublished results), but we hypothesized that design modifications would allow higher activity to be achieved in human cells.

The trans-ribozyme molecular design strategies address challenges in the cleavage activity and hybridization efficiency in adapting the unimolecular cis-acting system to the bimolecular trans-acting system. However, in implementing a trans-ribozyme in a cellular system the next level of design must address the stability and localization of the trans-acting molecule, two critical factors in the efficacy of trans-acting RNA regulatory systems. Preliminary experiments previously conducted in the Smolke laboratory have demonstrated that these two factors limit the regulatory activity of trans-ribozymes in yeast cells (Kate Galloway, unpublished results). However, it is likely that differences in the time scales of RNA transcription, processing, trafficking, and degradation may allow trans-ribozymes to function more effectively in human cells than in yeast.

Three trans-ribozymes were designed and tested in human cells (Figure 2.2). Two of the trans-ribozymes are derived from previously studied³⁴ hammerhead ribozymes: sTRSV and peach latent mosaic viroid (PLMVd). The third trans-ribozyme is based on the core of PLMVd but has modified stems designed to target a sequence within the coding region of a yeast-enhanced green fluorescent protein (yEGFP). The trans-ribozymes are flanked immediately upstream and downstream by small hairpins, intended to insulate the trans-ribozyme sequence from the surrounding transcript and prevent

intramolecular binding of the targeting arms, which must remain single-stranded in order to bind to the target sequence. Each trans-ribozyme is coupled with a unique targeting sequence, such that hybridization of the targeting arms reconstitutes stems I and III, forming a catalytically active ribozyme. The targeting sequences are placed within the 3' UTR of EGFP such that cleavage can be detected by monitoring fluorescence levels. Additionally, the targeting sequences are placed in multiple copies within the 3' UTR in order to examine the regulatory activity of the trans-ribozymes as a function of the number of target sites.

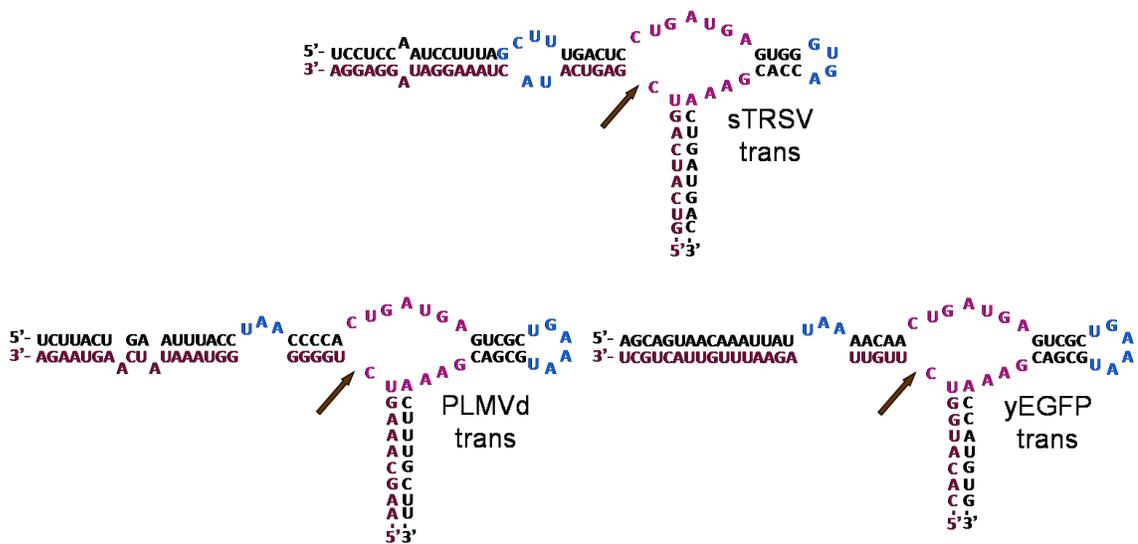


Figure 2.2. Structures of trans-ribozymes bound to target sequences. Coloring is the same as in Figure 2.1.

Characterization of initial trans-ribozyme designs in a human cell line

The trans-ribozymes are expressed from either a cytomegalovirus (CMV) RNA polymerase II (Pol II) or a U6 RNA polymerase III (Pol III) promoter. Pol II promoters

generally control the synthesis of mRNAs, which are capped on their 5' ends with 7-methylguanosine and polyadenylated on their 3' ends. The 5' cap and poly(A) tail associate with one another through a complex of proteins, thereby forming a circular messenger ribonucleoprotein (mRNP) complex that exhibits greater resistance to decapping enzymes and thus increased stability. In contrast, Pol III promoters generally control the synthesis of small non-coding RNAs, such as microRNAs (miRNAs) and small interfering RNAs (siRNAs), and do not have a 5' cap or poly(A) tail. The trans-ribozyme gene is assembled on a plasmid containing the fluorescent reporter gene DsRed-Express, which enables gating for cells that have been transfected with the plasmid (Figure 2.3).

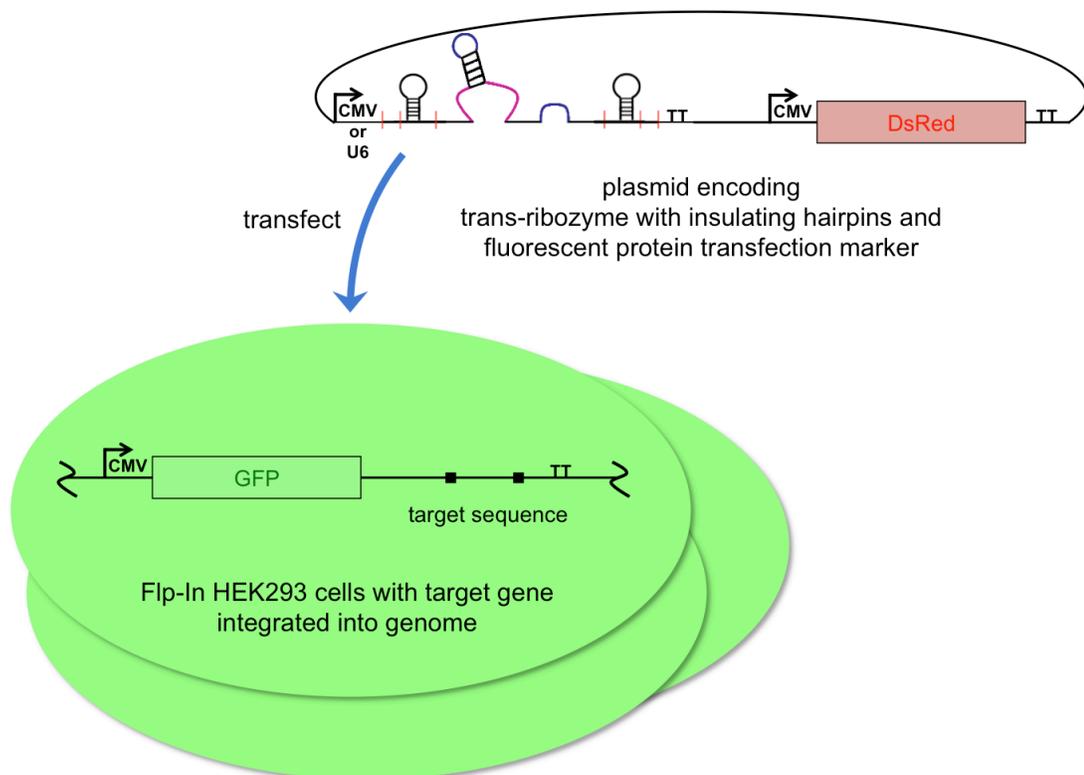


Figure 2.3. Trans-ribozyme and target gene characterization system. The trans-ribozyme is inserted between two insulating hairpins in a multiple cloning site (red lines). The resulting plasmid is transfected into cells with EGFP and the target sequence stably integrated into the genome.

EGFP and the target sequence(s) in its 3' UTR are stably integrated into the genome of human embryonic kidney 293 (HEK293) cells using the Flp-In system to generate isogenic stable cell lines (Figure 2.3). The gene is inserted into a plasmid backbone containing a Flp Recombination Target (FRT) site, thus allowing stable integration through genetic recombination in cell lines that have been engineered to contain a single copy of the FRT site in their genome. Integrating the fluorescent reporter gene in this way enables effective modeling of the targeting of endogenous transcripts by exogenous trans-ribozymes.

To quantify trans-ribozyme regulatory activity, stable cell lines expressing GFP with target sequence(s) were transiently transfected with a plasmid encoding a trans-ribozyme (Figure 2.3). GFP fluorescence was measured using flow cytometry, gating for transfected cells so that only cells harboring the plasmid encoding a trans-ribozyme were analyzed. Decreased GFP fluorescence is expected to correlate with increased regulatory activity. Analysis of the fluorescence of stable cell lines demonstrates that GFP constructs containing one copy of the target sequence are expressed at a higher level than GFP constructs containing multiple (2x or 4x) copies (Figure 2.4). These results indicate that the presence of target sequences in the 3' UTR may have some nonspecific effect on the expression of the target gene, potentially through transcript destabilization or

translational efficiency. However, even with these nonspecific effects, the data clearly indicate that none of the trans-ribozyme designs are able to downregulate expression of the target gene in this assay (Figure 2.4).

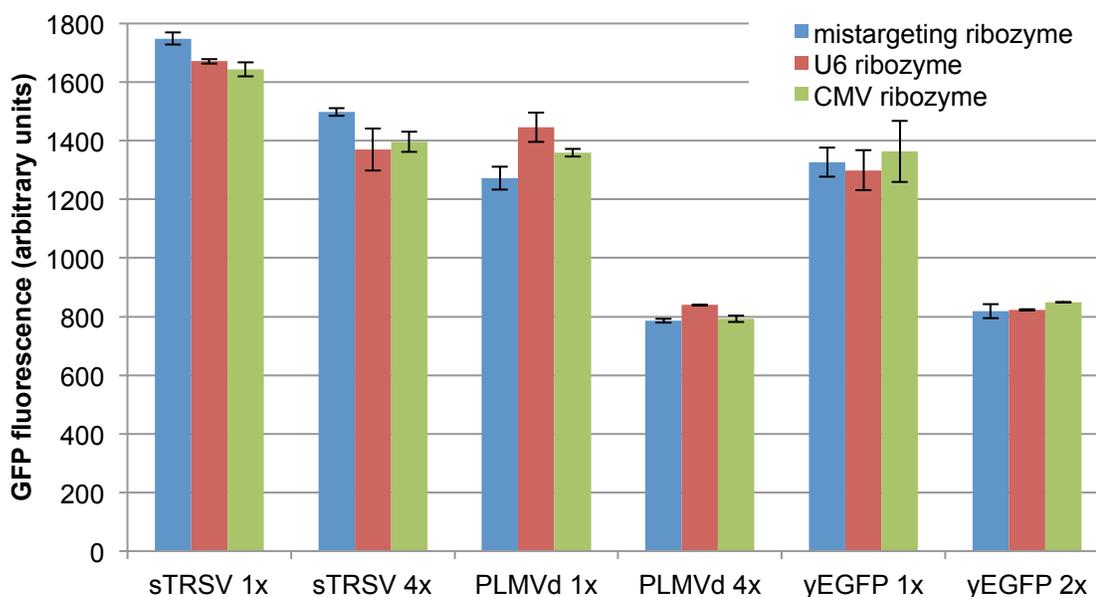


Figure 2.4. Activity of trans-ribozymes. GFP fluorescence levels are reported for stably integrated constructs encoding one or multiple copies of trans-ribozyme target sequences transfected with constructs encoding trans-ribozymes. Mistargeting trans-ribozymes that do not bind to the target sequence are included for comparison. Reported values are geometric mean \pm s.d. from biological duplicates.

There are two possible explanations for why the trans-ribozymes do not exhibit gene silencing activity. One possibility is that the two strands may not properly anneal *in vivo* to form a catalytically active ribozyme. A second possibility is that although the two strands properly anneal, the ribozyme as formed does not cleave at a sufficient rate to downregulate gene expression. It has previously been demonstrated that ribozyme

cleavage rate is correlated with *in vivo* gene knockdown³⁷, and specifically that if the cleavage rate is too low then gene regulation will not be observed.

Cis-ribozymes as a model for optimizing *in vivo* cleavage activity

To investigate whether the cleavage rate of the formed trans-ribozyme would be sufficient to observe knockdown *in vivo*, we designed Type I cis-ribozymes based on the sTRSV trans-ribozyme. Such designs remove the variable of whether the two strands can properly anneal in the cell and allow investigation of gene knockdown through ribozyme cleavage. K was formed by adding a GUUG tetraloop to the end of Stem III of the sTRSV trans-ribozyme (Figure 2.5), covalently joining the two strands into one. W is based on K but more closely resembles wild-type sTRSV, and Y even more so; W has the stem III sequence of sTRSV and Y is identical to W but with the loop I sequence reverted to that of wild-type sTRSV. CU is identical to Y except that the distal portion of stem I is integrated into a different position in bulge I. CU LsIII is identical to CU but with stem III extended by four base pairs, and CU LsIII inversion is identical to CU LsIII but with a stem III A-U pair changed to U-A. CK LsI, CK LsIII, and CK LsIV are all identical to CU LsIII but with the sequences of stem I, stem III, or both, respectively, from K. U LsIII is identical to CU LsIII except that the distal portion of stem I is integrated into a different position in bulge I. Finally, HHe-PLMVd is adapted from a previously described trans-ribozyme⁴, and 3-way AA and 3-way AAA are based on CU LsIII but include a three-way junction with an additional helix in stem I³⁸.

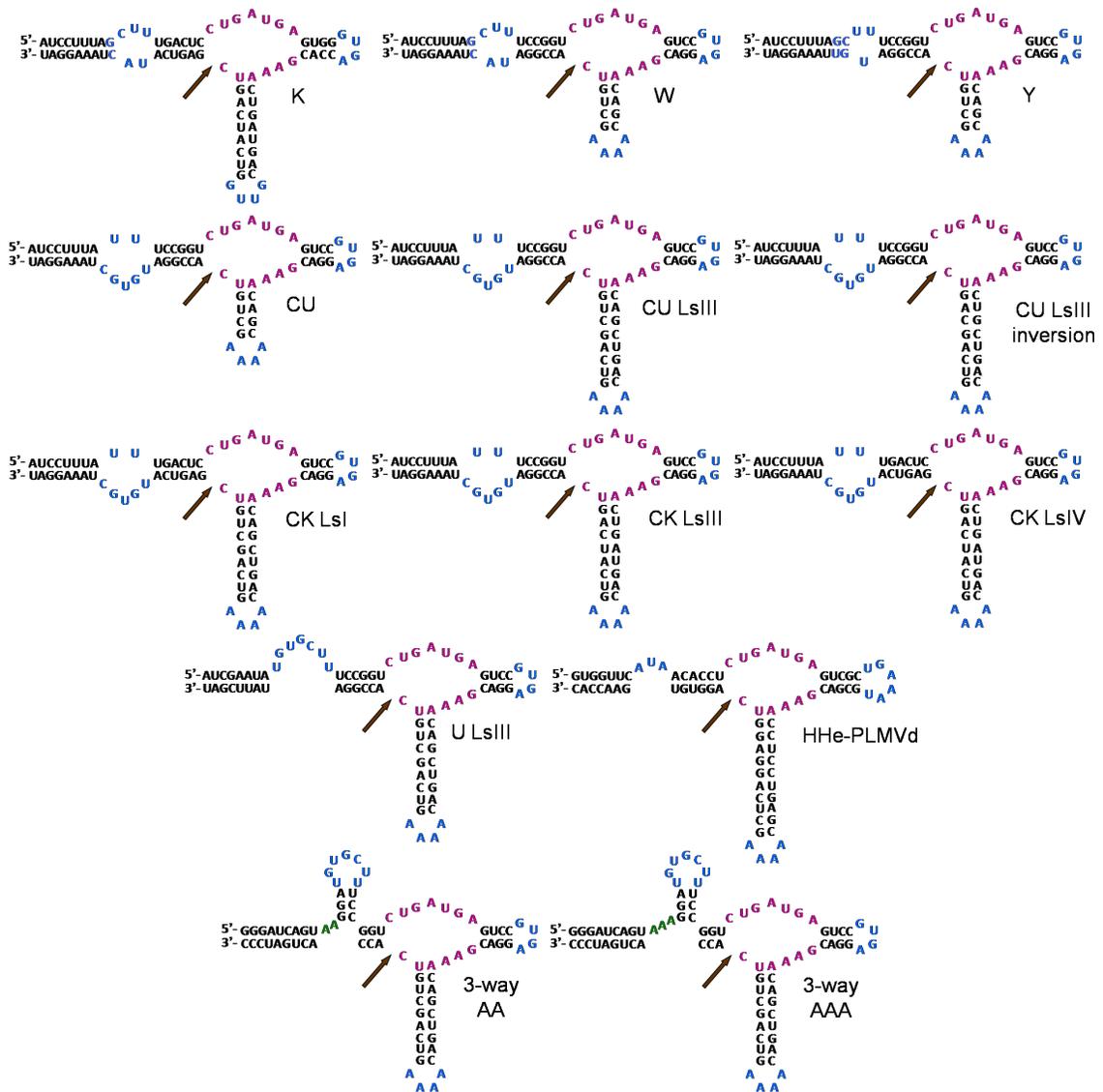


Figure 2.5. Structures of cis-ribozymes used to model trans-ribozyme activity. Stems are shown in black, the catalytic core is shown in magenta, and loops and bulges are shown in blue.

Characterization of Type I cis-ribozymes in a human cell line

Each Type I cis-ribozyme was placed in the 3' UTR of EGFP and gene regulation activity was measured in transient transfection assays by flow cytometry. K exhibited very little activity, with GFP fluorescence 80% of the non-cleaving control (Figure 2.6), indicating that the sTRSV trans-ribozyme would likely not be able to silence its target *in vivo*. The sequence of W is more closely related to wild-type sTRSV and Y even more so, and the activity of these designs reflects this. The alteration of the stem I integration point in CU leads to greater activity, and the extension of stem III in CU LsIII leads to a level of activity approaching that of sTRSV, with 8% expression compared to non-cleaving control. Inversion of the A-U base pair had a small detrimental effect on activity, while the three CK designs exhibited better activity the more similar they were to CU LsIII. Finally, U LsIII and HHe-PLMVd exhibited high levels of regulatory activity, while the 3-way designs showed little activity. These results suggest that the sequences of the ribozyme stems are not as flexible as previously reported^{2,35,39-41}.

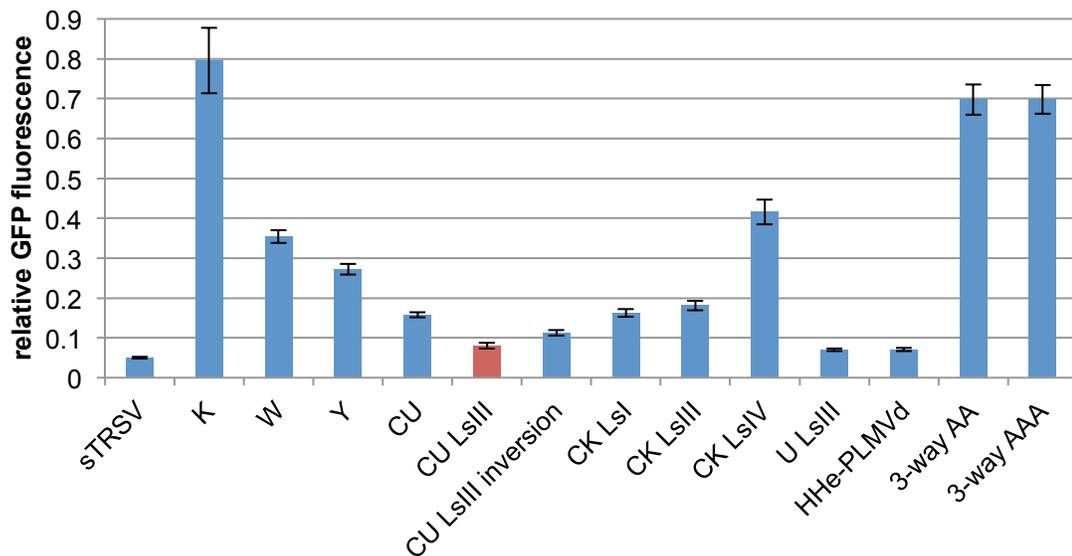


Figure 2.6. Activity of type I cis-ribozymes. Type I cis-ribozymes model the activity of trans-ribozymes. Relative GFP fluorescence levels are reported for transiently transfected constructs encoding ribozyme switch sequences. Reported values are geometric mean \pm s.d. from biological duplicates or triplicates and normalized to the non-cleaving sTRSVctrl.

Development of an improved trans-ribozyme

Based on its high level of gene-regulatory activity CU LsIII was chosen as the basis for a new trans-ribozyme design (Figure 2.7A). *In vitro* cleavage assays were performed to confirm the binding activity of the CU LsIII trans-ribozyme. For these experiments, the trans-ribozyme and target strands were synthesized using *in vitro* transcription, purified, and denatured and renatured separately. The RNA strands were then incubated together in a buffer representative of physiological conditions (500 μ M MgCl₂, 100mM NaCl, 50mM Tris-HCl (pH 7.5) at 37°C). Under these assay conditions, CU LsIII exhibited a cleavage rate of $\sim 0.3 \text{ min}^{-1}$ (Figure 2.7B), which is comparable to the cleavage rates of other ribozyme switches successfully used to regulate gene expression in yeast and mammalian cells^{25,26}. The results suggest that the CU LsIII trans-ribozyme is capable of binding and cleaving its target and should be capable of doing so *in vivo* at a rate sufficient for controlling gene expression levels.

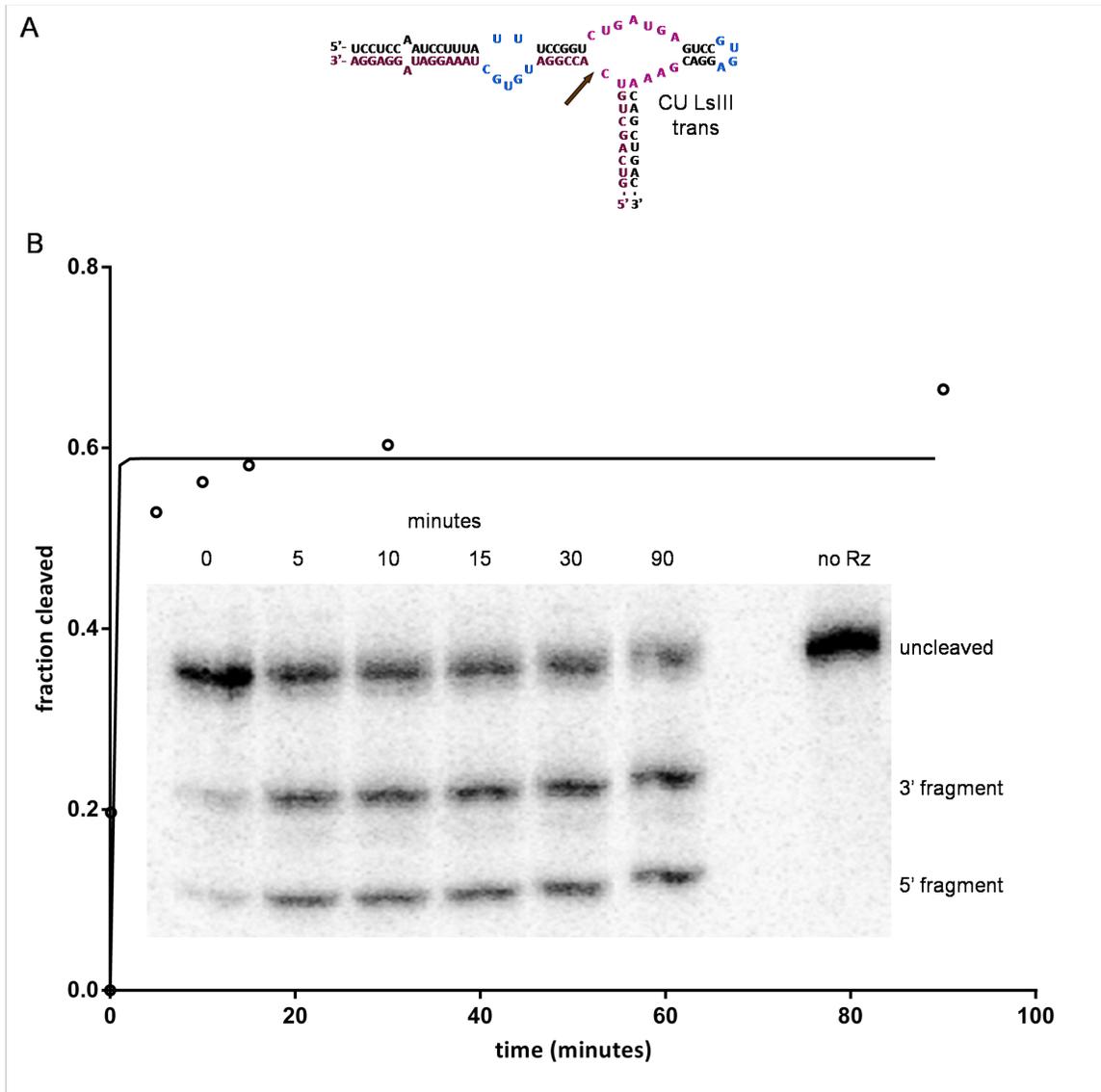


Figure 2.7. The improved trans-ribozyme. **(A)** Structure of the optimized CU LsIII trans-ribozyme. **(B)** *In vitro* cleavage activity of the CU LsIII trans-ribozyme. Cleavage of the internally radiolabeled target strand is monitored over time with PAGE, allowing calculation of the cleavage rate.

Incorporation of ancillary elements in the trans-ribozyme transcript to direct stability, structure, processing, and localization

There are a number of reasons why the trans-ribozymes described above may not be able to effectively cleave their target transcripts in human cells. The main obstacles are likely the stability of the trans-ribozyme strand in the cellular environment and the ability of this strand to bind to its target strand in the time scale of its lifetime. These issues are related, in that the less time required for the trans-ribozyme strand to bind to its target the less time it needs to exist in the cell, and the higher the stability of the trans-ribozyme strand the more time it will have to bind to its target. To address these issues, we developed a variety of expression constructs incorporating ancillary genetic elements into the sequence context of the CU LsIII trans-ribozyme.

When expressed from the CMV Pol II promoter, the trans-ribozyme is part of a longer transcript. Since the trans-ribozyme may interact with other parts of the transcript in a way that disrupts binding and cleaving of the target strand, we designed a construct containing cis-ribozymes immediately upstream and downstream of the trans-ribozyme and its insulating hairpins (Figure 2.8A). This construct was intended to function by cleaving the trans-ribozyme out of the transcript, potentially making the trans-ribozyme more accessible for binding to the target strand. However, the trans-ribozyme strand may be highly unstable once excised from the rest of the mRNP, so we also designed a construct containing large hairpins internal to the cis-ribozymes (Figure 2.8B). Following cis-ribozyme cleavage these large hairpins are expected to stabilize the 5' and 3' ends of the excised transcript, protecting the trans-ribozyme strand from RNA exonuclease activity. These large hairpins were also tested in constructs without cis-

ribozymes, in both the CMV Pol II and U6 Pol III promoter expression systems (Figure 2.8C).

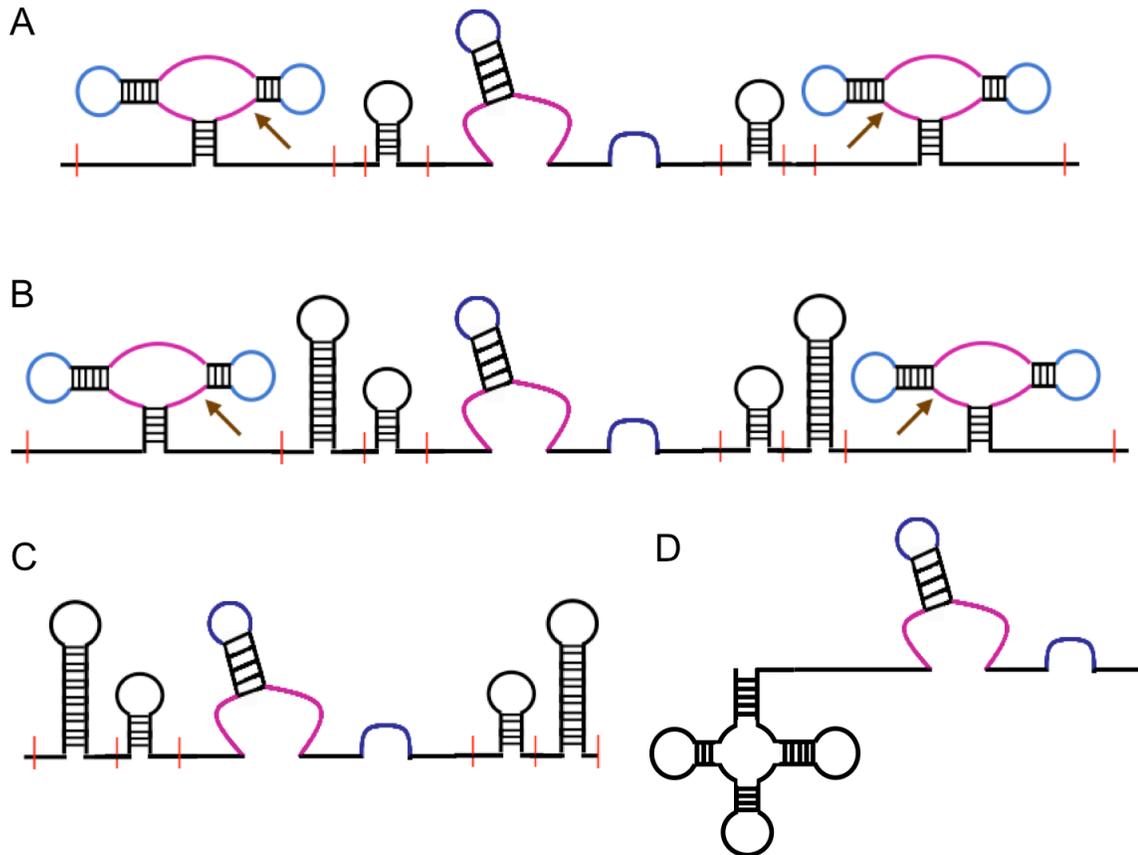


Figure 2.8. Trans-ribozyme ancillary elements. **(A)** Cis-ribozymes cleave the trans-ribozyme (with its insulating hairpins) out of the larger transcript. **(B)** Large hairpins stabilize the ends of the trans-ribozyme strand after excision. **(C)** Large hairpins stabilize Pol II and Pol III trans-ribozyme transcripts. **(D)** tRNA^{Val} stabilizes the trans-ribozyme transcript and localizes it to the cytoplasm.

Alternatively, we inserted tRNA^{Val} immediately upstream of the trans-ribozyme (Figure 2.8D), adapting work from Koseki and colleagues²⁹. They demonstrated that

their chimeric tRNA trans-ribozymes were highly stable in human cells and localized to the cytoplasm, and were able to cleave HIV-1 RNA *in vivo*³¹. We attempted to reproduce their work in our experimental system, using their HIV-targeting Rz2, as well as replacing the HIV trans-ribozyme with CU LsIII.

We assayed the ancillary elements with flow cytometry using transient transfections of stable lines as described above (Figure 2.3). None of the ancillary elements conferred activity on the CU LsIII trans-ribozyme (Figure 2.9). Additionally, we were unable to reproduce the activity of the HIV tRNA trans-ribozyme reported by Koseki and colleagues. We hypothesized that the stability of the GFP reporter used in our studies might be too high, such that significant protein levels remain even when the associated mRNA is cleaved by trans-ribozymes, masking the knockdown effect. To address this possibility, we replaced GFP with destabilized enhanced GFP (d2EGFP)⁴², which has a much shorter half-life than its parent. However, this modification to the experimental system did not result in detectable trans-ribozyme activity.

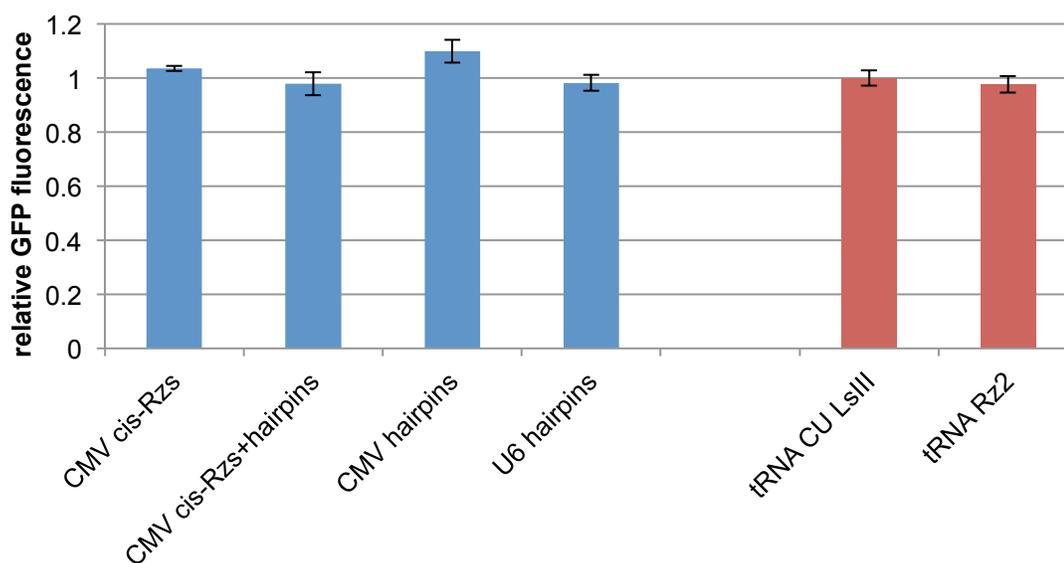


Figure 2.9. Activity of trans-ribozymes with ancillary elements. GFP fluorescence levels are reported for stably integrated constructs encoding trans-ribozyme target sequences transfected with constructs encoding trans-ribozymes with ancillary elements. Reported values are geometric mean \pm s.d. from biological duplicates and normalized to non-cleaving control trans-ribozymes.

Discussion

We attempted to develop a trans-acting version of the cis-acting ribozyme switch platform previously developed in the Smolke laboratory²⁴. Our trans-ribozyme architecture benefits from the previous engineering efforts directed to the cis-acting platform, in particular the design principles for aptamer integration with the ribozyme to build functional ligand-responsive gene-regulatory devices. The trans-ribozyme switch platform should have the added advantage of being able to regulate endogenous gene targets in trans in response to specified molecular inputs. However, there are additional requirements for such a trans-acting RNA device to function properly. Specifically, the functional RNA must be expressed in the cell at an appropriate concentration and localized appropriately, the binding site on the mRNA target must be accessible, and once annealed the duplex must fold into a catalytically active conformation.

Since we did not observe gene-regulatory activity from our initial trans-acting ribozyme designs (Figure 2.4), we first examined the ability of our designs to exhibit cleavage activity when annealed to the target sequence. Specifically, we constructed and

characterized cis-ribozyme versions of our trans-ribozyme designs, which were more likely to fold into the desired conformation based on the unimolecular context. Indeed, the cis version of the sTRSV trans-ribozyme exhibited minimal regulatory activity (Figure 2.6). We investigated the effects of stem length, stem and loop sequence, and loop I integration position on *in vivo* activity. We found that deviation from the sTRSV sequence in the ribozyme stems was detrimental to ribozyme function (Figure 2.6). This was surprising given that the stem sequence has generally been considered to be mutable due to its sequence diversity among natural hammerhead ribozymes, unlike the highly conserved catalytic core^{35,39-41}. These investigations led to the design of a new trans-ribozyme with a high degree of sequence similarity to sTRSV, which we used for all subsequent device optimization.

Following development of the optimized trans-ribozyme design, we further explored modifications to the design of the expression system that would support a high level of expression of the trans-ribozyme transcript. In particular, the transcription rate of the trans-ribozyme expression system was set to a high level by testing two strong promoters that act through different mechanisms, the CMV (Pol II) and U6 (Pol III) promoters. We further introduced design elements to reduce the degradation rate of the trans-ribozyme transcript by incorporating large hairpins on the 5' and 3' ends of the Pol III transcript and the unprotected portion of the Pol II transcript following excision from the mRNP mediated by cis-ribozymes. We also used a chimeric tRNA trans-ribozyme expression platform, which had previously been demonstrated to support gene regulation from a trans-ribozyme and shown to have a long half-life *in vivo*²⁹.

In addition to being present in cells at a sufficiently high concentration, a functional trans-ribozyme must also be localized to the same subcellular location as its target to be able to bind and cleave. Co-localization of the trans-ribozyme and target strands has been shown to be important for activity³, and we hypothesized that the chimeric tRNA trans-ribozymes would be transported to the cytoplasm, increasing their local concentration in the vicinity of their target mRNA and thereby improving hybridization efficiency. We incorporated hairpins immediately upstream and downstream of the trans-ribozyme to insulate it from the surrounding transcript, attempting to minimize misfolding that would occlude the targeting arms. We demonstrated with *in vitro* cleavage assays that the trans-ribozyme is capable of annealing with and cleaving its target sequence under physiological conditions (Figure 2.7).

Despite optimization of cleavage activity in model cis-ribozymes and incorporation of design elements to increase trans-ribozyme stability *in vivo*, we were unable to demonstrate trans-ribozyme-mediated gene-regulatory activity (Figure 2.9). We demonstrated with model cis-ribozymes that cleavage activity was sufficiently high to produce a large amount of gene knockdown, and we showed that the *in vitro* cleavage activity of our improved trans-ribozyme was comparable to that of previously characterized *in vivo* functional cis-ribozymes in yeast and mammalian systems. Taken together these results suggest that our trans-ribozyme did not function *in vivo* due to issues with transcript levels and co-localization with the target strand. Our efforts to improve trans-ribozyme stability and localization did not resolve these issues.

One possible area for further investigation is target site accessibility, which we did not address with any of our designs. It has been shown that in a typical mRNA many target sites will be inaccessible due to secondary and tertiary structure^{33,35}. Optimization of target site location within the target mRNA strand could lead to functional trans-ribozymes. However, successful regulation of an endogenous gene could require extensive screening of many trans-ribozymes targeting different target sites. Another possible strategy is employing RNA localization elements to target trans-ribozyme transcripts to the specific subcellular location of the mRNA target^{43,44}, increasing the local effective concentration and increasing the likelihood of hybridization between the two strands.

The finding that the sequence of the ribozyme stems is less flexible than expected coupled with the issue of target site accessibility severely limits the capability of trans-ribozymes to target endogenous genes. Furthermore, the independence of trans-ribozymes from cell-specific machinery makes them vulnerable to degradation, while other methods for regulating endogenous genes, such as miRNA⁴⁵ and clustered regularly interspaced palindromic repeats interference (CRISPRi)⁴⁶, benefit from protein complexes that protect the RNA and facilitate interaction with the target strand. Taken together, the limitations of the trans-ribozyme platform present a significant challenge to the regulation of endogenous genes, while other RNA-based platforms are more effective and promising. Ligand-responsive miRNAs have previously been demonstrated⁴⁵, and although allosteric regulation has not yet been demonstrated for CRISPRi, such capability may soon be realized. These platforms may therefore be better poised to provide programmable ligand-responsive regulation of endogenous genes.

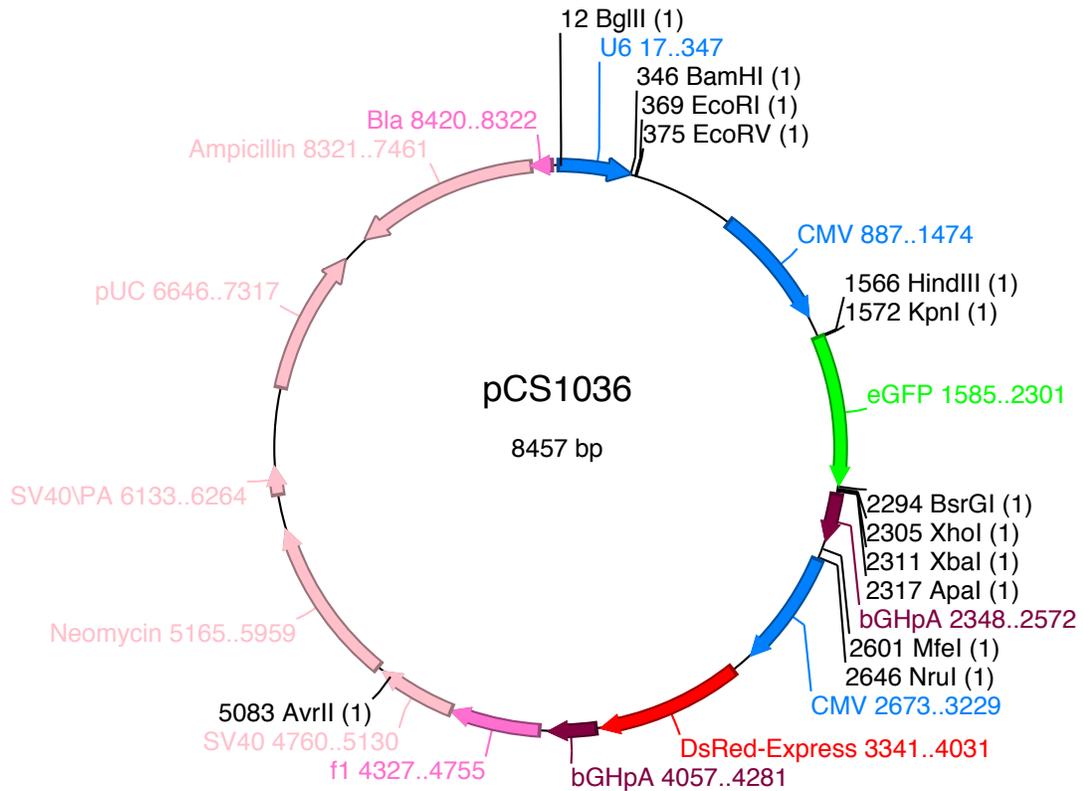
Methods

Plasmid construction

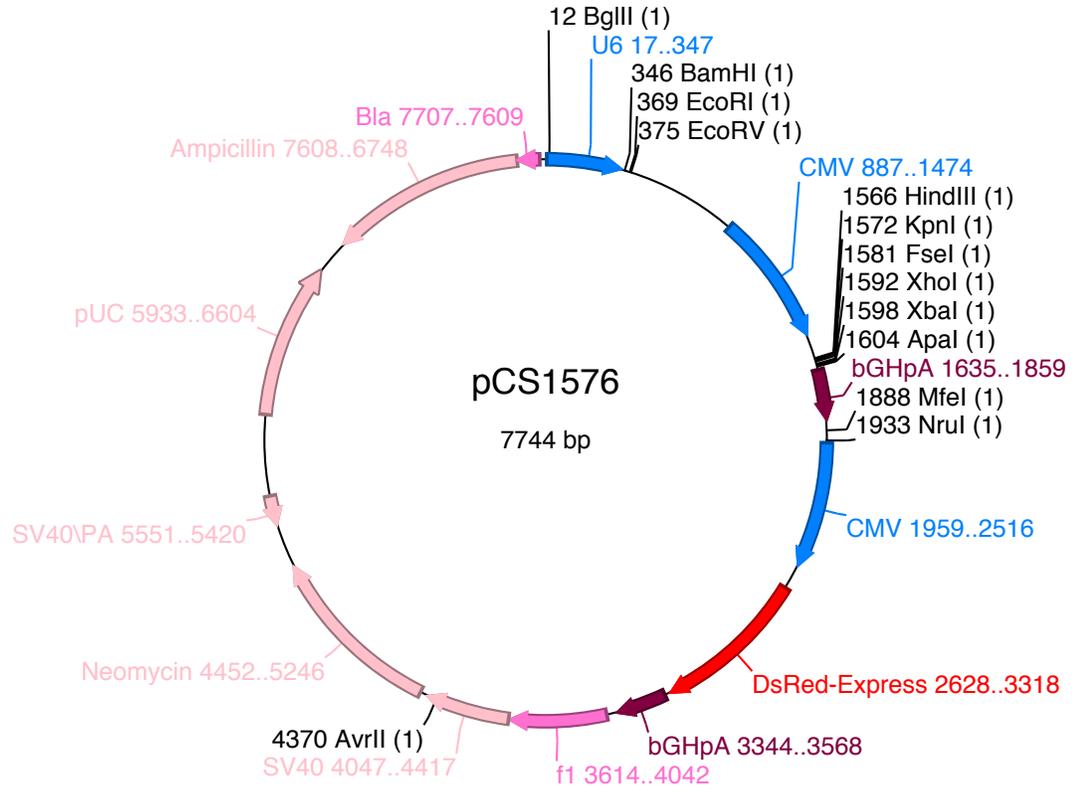
All plasmids were constructed using standard molecular biology techniques. Oligonucleotides were synthesized by Integrated DNA Technologies and the Stanford Protein and Nucleic Acid Facility. Cloning enzymes, including restriction enzymes and T4 DNA ligase, were obtained from New England Biolabs. Ligation products were electroporated into *Escherichia coli* DH10B (Life Technologies) using a GenePulser XP (Bio-Rad Laboratories) system using standard methods. Clones were screened using colony polymerase chain reaction (PCR) and verified by sequencing (Laragen Inc. and Elim Biopharmaceuticals). 15% glycerol stocks were made from *E. coli* in logarithmic growth phase and stored at -80°C .

A standardized cloning method was developed to facilitate insertion of trans-ribozymes into various sequence contexts. The DNA fragment insertFseI was inserted into pCS1036 (courtesy Yvonne Chen) (derived from pcDNA3.1(+) (Life Technologies)) between the restriction sites KpnI/XhoI to form pCS1576 (Figure 2.10), which contained a U6 and a CMV promoter for expressing trans-ribozymes and DsRed-Express as a transfection control. Cassettes containing restriction sites, a terminator (U6 only), and small hairpins designed to prevent intramolecular binding of the trans-ribozyme targeting arms (U6 trans-ribozyme cassette and CMV trans-ribozyme cassette) were inserted downstream of the U6 (between BamHI/EcoRI) and CMV (between FseI/XhoI) promoters to form pCS1646 and pCS1662, respectively (Figure 2.10). Ancillary cis-ribozymes were inserted into pCS1662 between HindIII/KpnI and XbaI/ApaI to form pCS1955. Ancillary large stabilizing hairpins were inserted into pCS1646 between

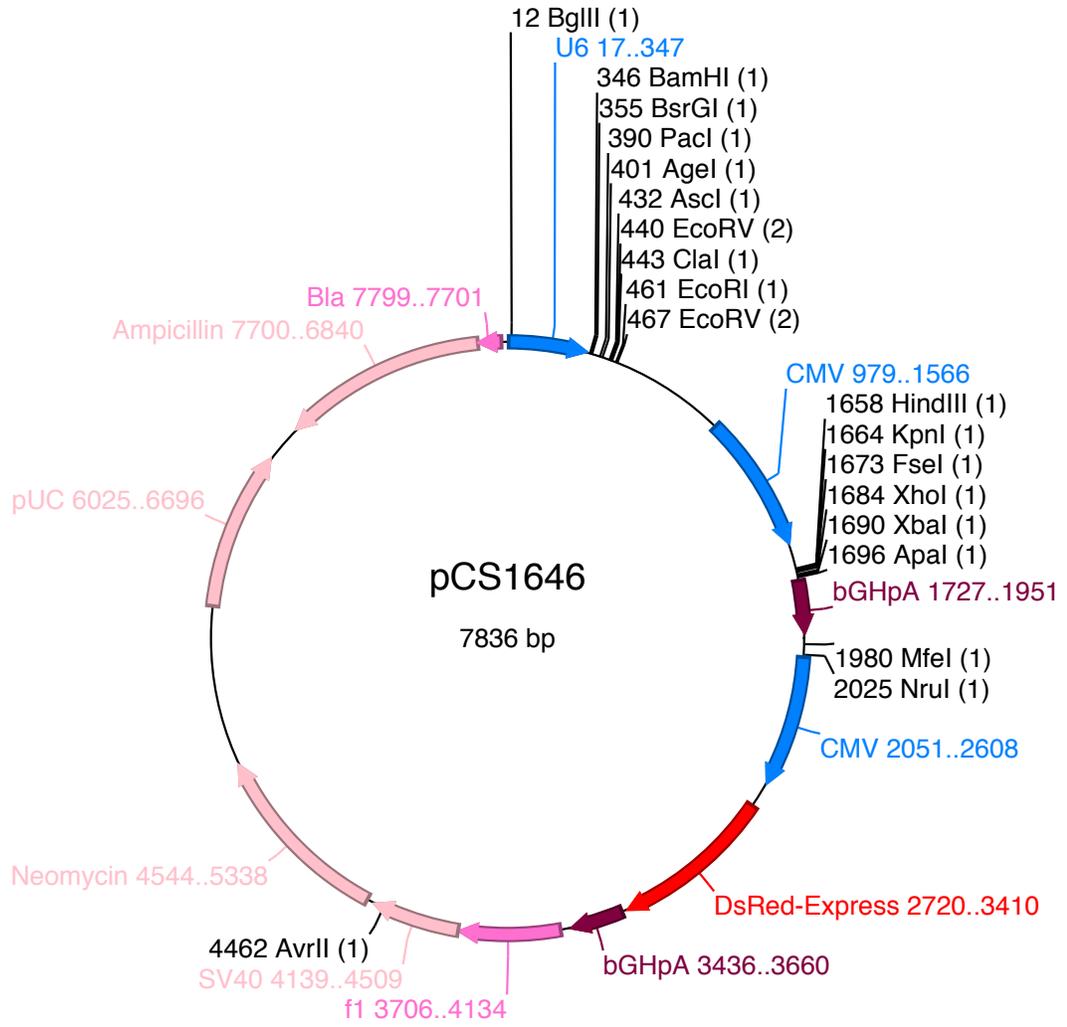
BamHI/BsrGI and AscI/ClaI and into pCS1662 between KpnI/FseI and XhoI/XbaI to form pCS1953 and pCS1954, respectively. Both ancillary cis-ribozymes and large stabilizing hairpins were inserted into pCS1662 (using the same restriction sites used to form pCS1955 and pCS1954) to form pCS1956. Trans-ribozymes were inserted into pCS1646, pCS1662, pCS1955, pCS1953, pCS1954, and pCS1956 between PacI/AgeI.

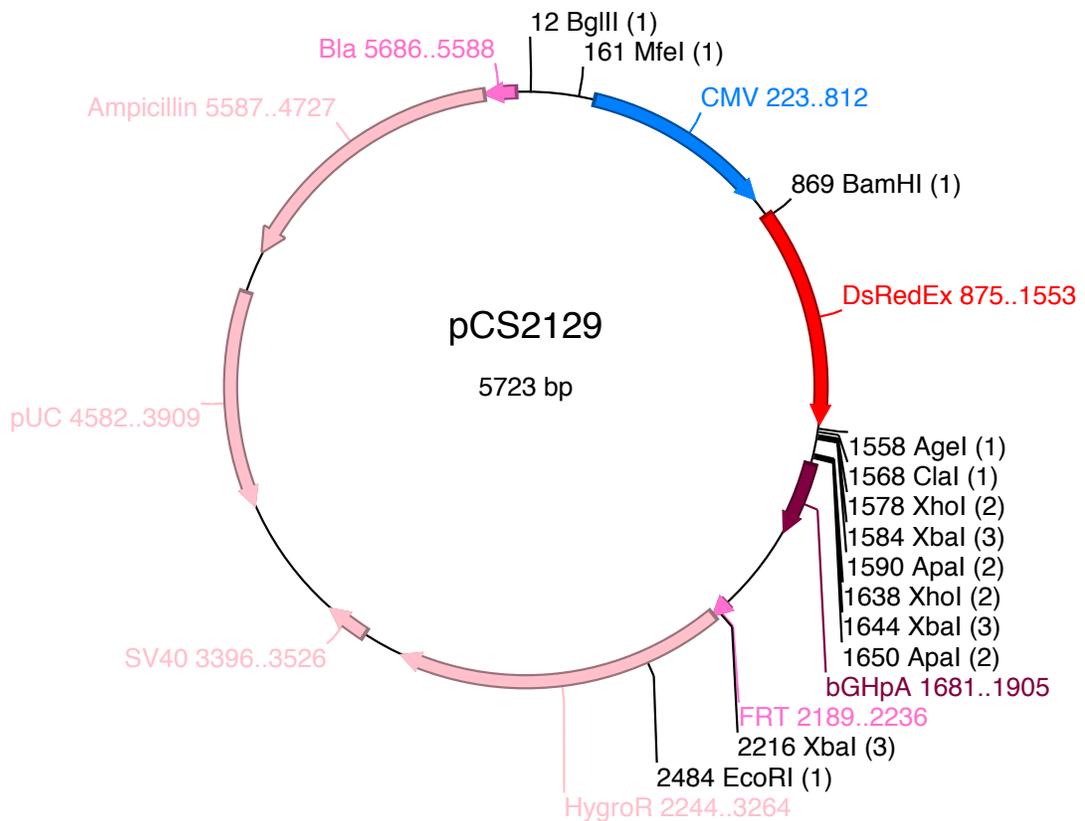
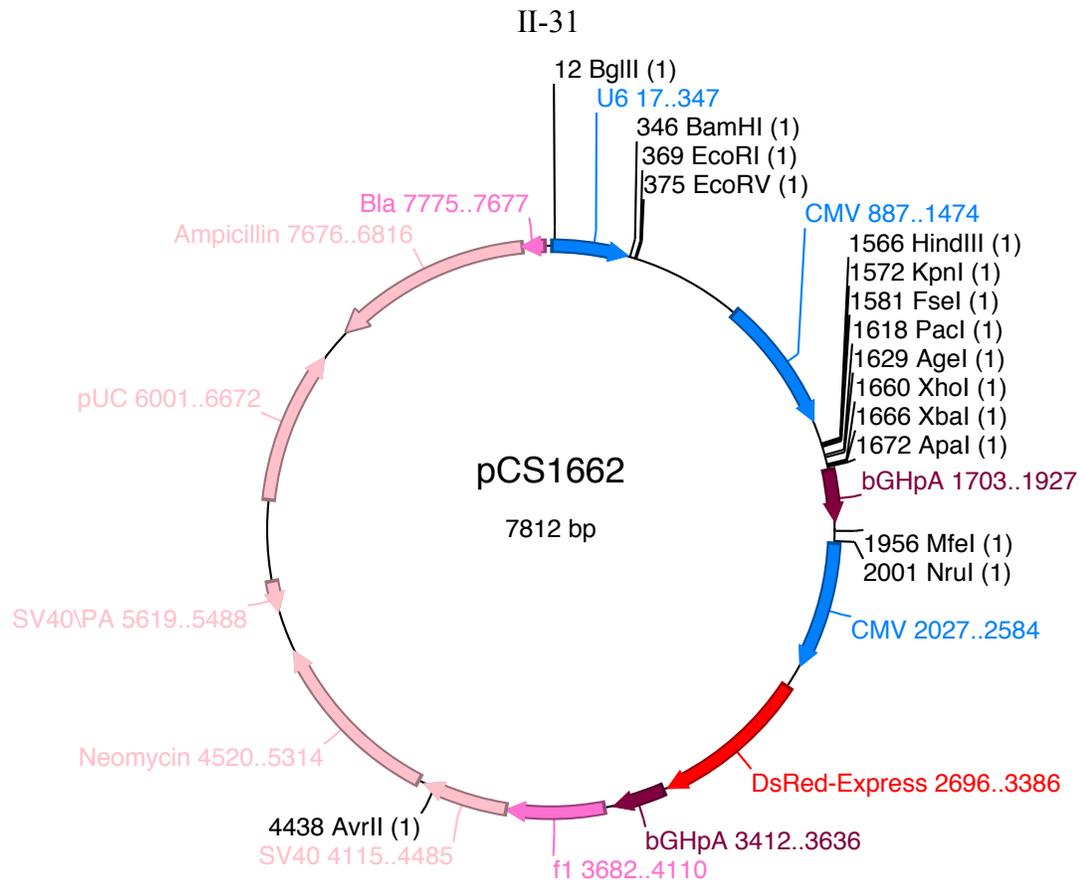


II-29



II-30





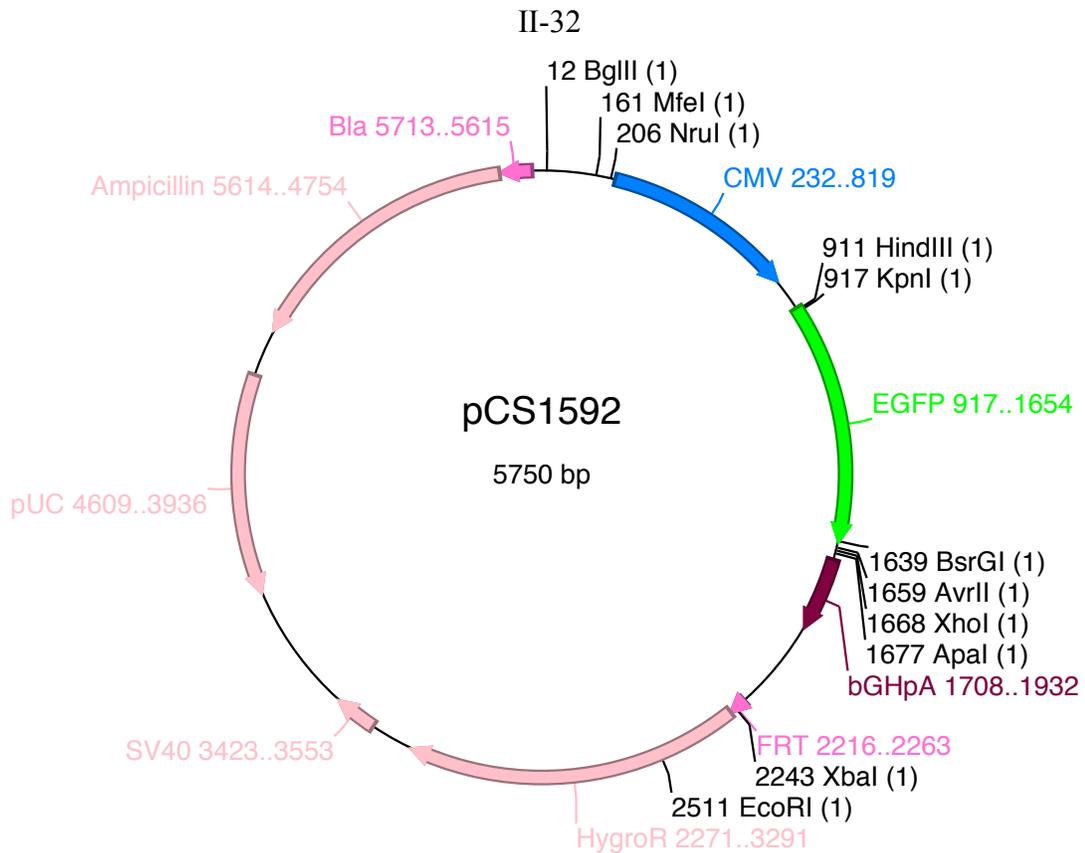


Figure 2.10. Plasmid maps.

The plasmid d2EGFP-Flp-In (courtesy Ryan Bloom) (derived from pcDNA5/FRT (Life Technologies)) was digested with NruI/EcoRV and blunt-end ligated to form pCS2129 (Figure 2.10), which contained DsRed-Express as a transfection control. Trans-ribozymes with tRNA^{Val} 5' and 3' sequences were inserted into pCS2129 between BglII/MfeI.

The DNA fragment insertAvrII was inserted into pCS1302 (courtesy Yvonne Chen) (derived from pcDNA5/FRT) between AvrII/ApaI to form pCS1592 (Figure 2.10), which contained a CMV promoter expressing EGFP. Trans-ribozyme target sequences in one or multiple copies with spacers were digested out of pCS1306 and pCS1642 (sTRSV), pCS1305 and pCS1495 (PLMVd), and pCS1492 and pCS1496 (yEGFP)

(courtesy Kate Galloway) and inserted into pCS1592 using AvrII/XhoI. CU LsIII and HIV target sequences were inserted into pCS1592 between AvrII/XhoI to form pCS1966 and pCS2603, respectively. The coding region of d2EGFP was PCR amplified from the plasmid d2EGFP-Flp-In using the primers d2eGFP HindIII 62 F and d2eGFP AvrII 62 R and inserted into pCS1966 between HindIII/AvrII to form pCS2147. The resulting plasmids were used to create isogenic stable cell lines through the Flp-In system (Life Technologies).

Type I ribozymes with spacers were inserted into pCS1036, which contained a CMV promoter expressing EGFP and DsRed-Express as a transfection control, between XhoI/ApaI (Figure 2.10).

Human cell culture

Flp-In HEK293 cells (Life Technologies) were cultured in 10 mL (10 cm dish) or 3 mL (6 cm dish) Dulbecco's Modified Eagle's medium (DMEM) (Life Technologies) supplemented with 10% fetal bovine serum (FBS) (Life Technologies) and 100 mg/L zeocin (Life Technologies) in a humidified incubator at 37°C and 5% CO₂. Cells were seeded at 2x10⁴ cells/mL and passaged regularly using 0.25% trypsin-EDTA (Life Technologies), with media replaced every 48–72 hours. Cells stably integrated with Flp-In constructs were cultured similarly, except the cell culture media were supplemented with 100 mg/L hygromycin B (Life Technologies) and no zeocin.

Stable cell line generation

Flp-In HEK293 cells were seeded at 1×10^5 cells/mL in 2 mL (6-well plate) DMEM with 10% FBS. 24 hours later the cells were cotransfected with a pcDNA5/FRT-derived plasmid and pOG44 (Life Technologies) in a 1:9 ratio using FuGENE 6 or FuGENE HD (Promega) according to the manufacturer's instructions. Typically DNA and FuGENE were incubated together in Opti-MEM in a 1:3:50 (g:L:L) ratio for approximately 1 hour, with 2 mL samples receiving 2 μ g of DNA. 24 hours after transfection the cells were resuspended using 0.25% trypsin-EDTA and DMEM with 10% FBS, and $\frac{1}{4}$ of the cells were used to seed 2 mL (6-well plate) DMEM with 10% FBS. 24 hours later the media were replaced with DMEM with 10% FBS and 200 mg/L hygromycin B. The media were replaced every 72–96 hours until macroscopic colonies were visible, usually after 10–14 days. Colonies were pooled together with 0.25% trypsin-EDTA and passaged into DMEM with 10% FBS and 100 mg/L hygromycin B. 10% dimethyl sulfoxide (DMSO) stocks were made from resuspended cells, cooled by 1 degree/minute to -80°C , then stored at -320°C .

Transient transfection

Flp-In HEK293 cells were seeded at 1×10^5 or 3×10^5 cells/mL in 500 μ L (24-well plate) DMEM with 10% FBS. 23–29 hours after seeding the cells were transfected with plasmid using FuGENE 6 or FuGENE HD according to the manufacturer's instructions. Typically DNA and FuGENE were incubated together in Opti-MEM in a 1:3:50 (g:L:L) ratio for approximately 1 hour, with 500 μ L samples receiving 500 ng of DNA.

Flow cytometry

40–48 after transfection fluorescence data were obtained by flow cytometry using the Quanta Cell Lab Flow Cytometer equipped with a 488 nm laser (Beckman Coulter). Viability was gated by side scatter and electronic volume, and viable cells were further gated for DsRed expression, which served as a transfection control. GFP and DsRed fluorescence was measured through 525/30 nm band-pass and 610 nm long-pass filters, respectively. Data were analyzed using FlowJo (Tree Star Inc.). Geometric mean values from biological replicates were reported with an error range of ± 1 standard deviation. Geometric mean fluorescence values were normalized to those of a control with no ribozyme or an inactive ribozyme.

***In vitro* cleavage assays**

The CU LsIII trans-ribozyme and its target strand were amplified by PCR from plasmids pCS1949 and pCS1966, respectively, using the primers CU HP T7 F and CU HP T7 R for the ribozyme strand and the primers Barcode T7 F and Barcode T7 R for the target strand. The forward primers added the T7 promoter sequence. Trans-ribozyme RNA was generated by *in vitro* transcription using 1 μg PCR product DNA as a template, with 40 mM Tris-HCl, pH 7.9, 16 mM MgCl_2 , 10 mM dithiothreitol (DTT), 2 mM spermidine, 3 mM rATP, rCTP, rGTP, and rUTP, 40 U RNaseOUT (Life Technologies), and 50 U T7 RNA polymerase (New England Biolabs) in 25 μL total volume and incubated at 37°C for approximately 2 hours. The transcription product was treated with 2 U of DNaseI (New England Biolabs) at 37°C for approximately 15 min and purified using the RNA Clean & Concentrator-25 kit (Zymo Research) according to the

manufacturer's instructions. Internally radiolabeled target strand RNA was generated by *in vitro* transcription using a similar method, except with rGTP reduced to 300 μ M and supplemented with 5 μ Ci [α - 32 P]rGTP.

Trans-ribozyme and radiolabeled target RNA were denatured separately by heating to 95°C in 50 mM Tris-HCl and 100 mM NaCl, then cooled by 1.2°C/minute to 37°C. Trans-ribozyme (1 μ M final concentration) was added to target (100 nM final concentration) and the reaction was initiated by adding MgCl₂ (500 μ M final concentration) and incubating at 37°C. Aliquots were removed and quenched with RNA stop/load buffer (95% formamide, 30 mM EDTA, 0.25% bromophenol blue, 0.25% xylene cyanol) on ice. Reaction products were heated to 95°C for 5 min, snap cooled on ice for 5 min, and separated by 12% (w/v) denaturing polyacrylamide gel electrophoresis (PAGE) with 8.3M urea. The 32 P radioactivity of cleaved and uncleaved bands was quantified by phosphorimager analysis using a Molecular Imager FX (Bio-Rad Laboratories) and Quantity One software (Bio-Rad Laboratories).

Acknowledgements

We thank Kate Galloway for design of the initial trans-ribozymes and providing plasmid constructs, and Ryan Bloom and Yvonne Chen for providing plasmid constructs. This work was supported by the National Institutes of Health (grant to Christina Smolke, RC1GM091298).

References

1. Guerrier-Takada, C. et al. Phenotypic conversion of drug-resistant bacteria to drug sensitivity. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 8468–8472 (1997).
2. Sarver, N. et al. Ribozymes as potential anti-HIV-1 therapeutic agents. *Science*. **247**, 1222–1225 (1990).
3. Gu, S. et al. Inhibition of infectious human immunodeficiency virus type 1 virions via lentiviral vector encoded short antisense RNAs. *Oligonucleotides* **16**, 287–295 (2006).
4. Carbonell, A. et al. Trans-cleaving hammerhead ribozymes with tertiary stabilizing motifs: in vitro and in vivo activity against a structured viroid RNA. *Nucleic Acids Res.* **39**, 2432–2444 (2011).
5. Hampel, A. The Hairpin Ribozyme: Discovery, Two-Dimensional Model, and Development for Gene Therapy. *Prog. Nucleic Acid Res. Mol. Biol.* **58**, 1–39 (1998).
6. Yu, M. et al. Intracellular immunization of human fetal cord blood stem/progenitor cells with a ribozyme against human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 699–703 (1995).
7. McCaffrey, A. P. et al. Inhibition of hepatitis B virus in mice by RNA interference. *Nat. Biotechnol.* **21**, 639–644 (2003).
8. Kuwabara, T. et al. A novel allosterically trans-activated ribozyme, the maxizyme, with exceptional specificity in vitro and in vivo. *Mol. Cell* **2**, 617–627 (1998).
9. Kato, Y. et al. Relationships between the activities in vitro and in vivo of various kinds of ribozyme and their intracellular localization in mammalian cells. *J. Biol. Chem.* **276**, 15378–15385 (2001).
10. Zhang, Z. et al. The ABCC4 gene is a promising target for pancreatic cancer therapy. *Gene* **491**, 194–199 (2012).
11. Zhou, W. et al. ShRNA silencing glycogen synthase kinase-3 beta inhibits tumor growth and angiogenesis in pancreatic cancer. *Cancer Lett.* **316**, 178–186 (2012).
12. Michienzi, A. et al. Ribozyme-mediated inhibition of HIV 1 suggests nucleolar trafficking of HIV-1 RNA. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8955–8960 (2000).
13. Li, M. et al. Inhibition of HIV-1 infection by lentiviral vectors expressing pol III-promoted anti-HIV RNAs. *Mol. Ther.* **8**, 196–206 (2003).

14. Wong-Staal, F. et al. A controlled, Phase 1 clinical trial to evaluate the safety and effects in HIV-1 infected humans of autologous lymphocytes transduced with a ribozyme that cleaves HIV-1 RNA. *Hum. Gene Ther.* **9**, 2407–2425 (1998).
15. Macpherson, J. L. et al. Long-term survival and concomitant gene expression of ribozyme-transduced CD4+ T-lymphocytes in HIV-infected patients. *J. Gene Med.* **7**, 552–564 (2005).
16. Mitsuyasu, R. T. et al. Phase 2 gene therapy trial of an anti-HIV ribozyme in autologous CD34+ cells. *Nat. Med.* **15**, 285–292 (2009).
17. Culler, S. J. et al. Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins. *Science.* **330**, 1251–1255 (2010).
18. Saito, H. et al. Synthetic human cell fate regulation by protein-driven RNA switches. *Nat. Commun.* **2**, 160 (2011).
19. Chang, A. L. et al. Synthetic RNA switches as a tool for temporal and spatial control over gene expression. *Curr. Opin. Biotechnol.* **23**, 679–688 (2012).
20. Liu, Y. et al. Synthesizing oncogenic signal-processing systems that function as both “signal counters” and “signal blockers” in cancer cells. *Mol. Biosyst.* **9**, 1909–1918 (2013).
21. Endo, K. et al. Quantitative and simultaneous translational control of distinct mammalian mRNAs. *Nucleic Acids Res.* **41**, e135 (2013).
22. Buskirk, A. et al. Engineering a ligand-dependent RNA transcriptional activator. *Chem. Biol.* **11**, 1157–1163 (2004).
23. Buzayan, J. M. et al. Nucleotide sequence and newly formed phosphodiester bond of spontaneously ligated satellite tobacco ring spot virus RNA. *Nucleic Acids Res.* **14**, 9729–9743 (1986).
24. Win, M. N. & Smolke, C. D. A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14283–14288 (2007).
25. Liang, J. C. et al. A high-throughput, quantitative cell-based screen for efficient tailoring of RNA device activity. *Nucleic Acids Res.* **40**, e154 (2012).
26. Wei, K. Y. et al. A yeast-based rapid prototype platform for gene control elements in mammalian cells. *Biotechnol. Bioeng.* **110**, 1201–1210 (2013).
27. Chen, Y. Y. et al. Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 8531–8536 (2010).

28. Saksmerprome, V. et al. Artificial tertiary motifs stabilize trans-cleaving hammerhead ribozymes under conditions of submillimolar divalent ions and high temperatures. *RNA* **10**, 1916–1924 (2004).
29. Koseki, S. et al. Factors governing the activity in vivo of ribozymes transcribed by RNA polymerase III. *J. Virol.* **73**, 1868–1877 (1999).
30. Bertrand, E. et al. The expression cassette determines the functional activity of ribozymes in mammalian cells by controlling their intracellular localization. *RNA* **3**, 75–88 (1997).
31. Kuwabara, T. et al. Significantly higher activity of a cytoplasmic hammerhead ribozyme than a corresponding nuclear counterpart: engineered tRNAs with an extended 3' end can be exported efficiently and specifically to the cytoplasm in mammalian cells. *Nucleic Acids Res.* **29**, 2780–2788 (2001).
32. Sano, M. et al. Novel method for selection of tRNA-driven ribozymes with enhanced stability in mammalian cells. *Antisense Nucleic Acid Drug Dev.* **12**, 341–352 (2002).
33. Warashina, M. et al. RNA-protein hybrid ribozymes that efficiently cleave any mRNA independently of the structure of the target RNA. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 5572–5577 (2001).
34. Weinberg, M. S. & Rossi, J. J. Comparative single-turnover kinetic analyses of trans-cleaving hammerhead ribozymes with naturally derived non-conserved sequence motifs. *FEBS Lett.* **579**, 1619–1624 (2005).
35. Birikh, K. R. et al. The structure, function and application of the hammerhead ribozyme. *Eur. J. Biochem.* **245**, 1–16 (1997).
36. Khvorova, A. et al. Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. *Nat. Struct. Biol.* **10**, 708–712 (2003).
37. Kennedy, A. B. et al. A versatile cis-blocking and trans-activation strategy for ribozyme characterization. *Nucleic Acids Res.* **41**, e41 (2013).
38. Wieland, M. et al. Expanded hammerhead ribozymes containing addressable three-way junctions. *RNA* **15**, 968–976 (2009).
39. Koizumi, M. et al. Cleavage of specific sites of RNA by designed ribozymes. *FEBS Lett.* **239**, 285–288 (1988).
40. Uhlenbeck, O. C. A small catalytic oligoribonucleotide. *Nature* **328**, 596–600 (1987).

41. Flores, R. et al. Hammerhead ribozyme structure and function in plant RNA replication. *Methods Enzymol.* **341**, 540–552 (2001).
42. Li, X. et al. Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter. *J. Biol. Chem.* **273**, 34970–34975 (1998).
43. Singer, R. H. RNA zipcodes for cytoplasmic addresses. *Curr. Biol.* **3**, 719–721 (1993).
44. Kislauskis, E. H. et al. Sequences responsible for intracellular localization of beta-actin messenger RNA also affect cell phenotype. *J. Cell Biol.* **127**, 441–451 (1994).
45. Beisel, C. L. et al. Design of small molecule-responsive microRNAs based on structural requirements for Drosha processing. *Nucleic Acids Res.* **39**, 2981–2994 (2011).
46. Qi, L. S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).

Supplementary Tables

Supplementary Table 2.1. Primer and oligonucleotide sequences.

Name	DNA sequence
insertFseI	AAAGGCCGGCCAAA
insertAvrII	AAACTCGAGAAA
sTRSV trans	ATCCTCCAATCCTTTAGCTTTGACTCCTGATGAGTGGGTGA CCACGAAACTGATGAC
sTRSV target sequence	GTCATCAGTCGAGTCATACTAAAGGATAGGAGGAAT
PLMVd trans	TCTTACTGAATTTACCTAACCCCACTGATGAGTCGCTGAAA TGCGACGAAACTTTGCTT
PLMVd target sequence	AAGCAAAGTCTGGGGGGTAAATATCAAGTAAGA
yEGFP trans	AGCAGTAACAAATTCTTAAAACAACCTGATGAGTCGCTGAA ATGCGACGAAACCATGTG
yEGFP target sequence	CACATGGTCTTGTTAGAATTTGTTACTGCT
CU LsIII trans	ATCCTCCAATCCTTTATTTCCGGTCTGATGAGTCCGTGAGG ACGAAACAGCTGAC
CU LsIII ctrl trans	ATCCTCCAATCCTTTATTTCCGGTGTACTGTGTCCGTGAGG ACCGAACAGCTGAC
CU LsIII target sequence	GTCAGCTGTCACCGGATGTGCTAAAGGATAGGAGGA
HIV trans	ACACAACACTGATGAGGACCGAAAGGTCCGAAACGGGCAC
HIV ctrl trans	ACACAACACTAATGAGGACCGAAAGGTCCGAAACGGGCAC
HIV target sequence	GTGCCCGTCTGTTGTGT
target sequence 5' spacer	TAAATCTAGGAAACAAA
target sequence 3' spacer	ATAAACAAACTCGATCCGCGAAAAAACCGCGGA
U6 trans-ribozyme cassette	GTCTGTACAGGTGTCTTCTTGAGCATGCTCAAGAGACATTA ATTAAACAACCGGTACGTCCATTACAAAGTAATGGACGTG GCGCGCCGATATCGATAAATTTTTTAAA
CMV trans-ribozyme cassette	GGTGTCTTCTTGAGCATGCTCAAGAGACATTAATTAACAA CCGGTACGTCCATTACAAAGTAATGGACGT
ancillary large hairpin U6 5'	GTGTCACCTGCAGTATTAGCAAATAATACATGCAAGTGAC
ancillary large hairpin U6 3' and CMV	GTCACCTGCAGTATTAGCAAATAATACATGCAAGTGAC

ancillary cis-ribozyme 5'	AAACAAAATCCTTTATTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCTGACAAAAGTCAGCTGTCACCGGATGTGCTA AAGGATAAAAAGA
ancillary cis-ribozyme 3'	AAACAAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCAAAAAGA
tRNA 5'	AGGACTAGTCTTTTAGGTCAAAAAGAAGAAGCTTTGTAACCGTTGGTTTCCGTAGTGTAGTGGTTATCACGTTTCGCCTAAC ACGCGAAAGGTCCCCGGTTCGAAACCGGGCACTACAA
tRNA 3'	GTCGGAAACGGTTTTTTTCTATCGCGTCGAC
sTRSVctrl	GCTGTCACCGGATGTGCTTTCCGGTACGTGAGGTCCGTGAGGACAGAACAGC
sTRSV	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGC
K	ATCCTTTAGCTTTGACTCCTGATGAGTGGGTGACCACGAAACTGATGACGTTGGTCATCAGTCGAGTCATACTAAAGGAT
W	ATCCTTTAGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCAAAAGCTGTCACCGGATACTAAAGGAT
Y	ATCCTTTAGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCAAAAGCTGTCACCGGATGTTAAAGGAT
CU	ATCCTTTATTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCAAAAGCTGTCACCGGATGTGCTAAAGGAT
CU LsIII	ATCCTTTATTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCTGACAAAAGTCAGCTGTCACCGGATGTGCTAAAGGAT
CU LsIII inversion	ATCCTTTATTTCCGGTCTGATGAGTCCGTGAGGACGAAACTGCTGACAAAAGTCAGCAGTCACCGGATGTGCTAAAGGAT
CK LsI	ATCCTTTATTTGACTCCTGATGAGTCCGTGAGGACGAAACAGCTGACAAAAGTCAGCTGTCGAGTCATGTGCTAAAGGAT
CK LsIII	ATCCTTTATTTCCGGTCTGATGAGTCCGTGAGGACGAAACTGATGACAAAAGTCATCAGTCACCGGATGTGCTAAAGGAT
CK LsIV	ATCCTTTATTTGACTCCTGATGAGTCCGTGAGGACGAAACTGATGACAAAAGTCATCAGTCGAGTCATGTGCTAAAGGAT
U LsIII	ATCGAATATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCTGACAAAAGTCAGCTGTCACCGGATATTCGAT
HHe-PLMVd	GTGGTTCATAACACCTCTGATGAGTCGCTGAAATGCGACGAAACCTCCTGAGCAAAAGCTCAGGAGGTCAGGTGTGAACCA C
3-way AA	GGGATCAGTAAGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCTGACAAAAGTCAGCTGTCACCACTGAT CCC
3-way AAA	GGGATCAGTAAAGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCTGACAAAAGTCAGCTGTCACCACTGAT TCCC
A 5' spacer	AAACAAACAAA
A 3' spacer	AAAAGAAAATAAAAATTTTTTGAA
B 5' spacer	AATAAATAAAA

B 3' spacer	CAAATAAACAAACACTC
d2eGFP HindIII 62 F	TAGAAGCTTATGGTGAGCAAGGGCGAGGAG
d2eGFP AvrII 62 R	AAGCCTAGGTTTTGCTACACATTGATCCTAGCAGAAGCACAGG
CU HP T7 F	TTCTAATACGACTCACTATAGGGTGTCTTCTTGAGCATGCTCAAGAGACATTAATTAATCCTC
CU HP T7 R	ACGTCCATTACTTTGTAATGGACGTACCGGTG
Barcode T7 F	TTCTAATACGACTCACTATAGGACAAGTAACTCGAAAAACCTAGGTAAATCTAGGAAACAAAGT
Barcode T7 R	CTTTCTCGAGTCCGCGGTTTTTTCGC

Supplementary Table 2.2. Plasmid constructs.

Plasmid	Description
pCS1592	pCS1302 with insertAvrII inserted between AvrII/ApaI
pCS1631	pCS1592 with sTRSV target sequence with spacers inserted between AvrII/XhoI
pCS1632	pCS1592 with sTRSV target sequence (4 copy) with spacers inserted between AvrII/XhoI
pCS1629	pCS1592 with PLMVd target sequence with spacers inserted between AvrII/XhoI
pCS1630	pCS1592 with PLMVd target sequence (4 copy) with spacers inserted between AvrII/XhoI
pCS1633	pCS1592 with yEGFP target sequence with spacers inserted between AvrII/XhoI
pCS1634	pCS1592 with yEGFP target sequence (4 copy) with spacers inserted between AvrII/XhoI
pCS1966	pCS1592 with CU LsIII target sequence with spacers inserted between AvrII/XhoI
pCS2063	pCS1592 with HIV target sequence with spacers inserted between AvrII/XhoI
pCS1576	pCS1036 with insertFseI inserted between KpnI/XhoI
pCS1646	pCS1576 with U6 trans-ribozyme cassette inserted between BamHI/EcoRI
pCS1662	pCS1576 with CMV trans-ribozyme cassette inserted between FseI/XhoI
pCS1655	pCS1646 with sTRSV trans inserted between PacI/AgeI
pCS1653	pCS1646 with PLMVd trans inserted between PacI/AgeI
pCS1657	pCS1646 with yEGFP trans inserted between PacI/AgeI
pCS1685	pCS1662 with sTRSV trans inserted between PacI/AgeI

pCS1683	pCS1662 with PLMVd trans inserted between PacI/AgeI
pCS1687	pCS1662 with yEGFP trans inserted between PacI/AgeI
pCS1949	pCS1662 with CU LsIII trans inserted between PacI/AgeI
pCS1955	pCS1662 with ancillary cis-ribozyme 5' inserted between HindIII/KpnI and ancillary cis-ribozyme 3' inserted between XbaI/ApaI
pCS1956	pCS1662 with ancillary cis-ribozyme 5' inserted between HindIII/KpnI, ancillary cis-ribozyme 3' inserted between XbaI/ApaI, and ancillary large hairpin U6 3' and CMV inserted between KpnI/FseI and XhoI/XbaI
pCS1954	pCS1662 with ancillary large hairpin U6 3' and CMV inserted between KpnI/FseI and XhoI/XbaI
pCS1953	pCS1646 with ancillary large hairpin U6 5' inserted between BamHI/BsrGI and ancillary large hairpin U6 3' and CMV inserted between AscI/ClaI
pCS2012	pCS1955 with CU LsIII trans inserted between PacI/AgeI
pCS2013	pCS1955 with CU LsIII ctrl trans inserted between PacI/AgeI
pCS2014	pCS1956 with CU LsIII trans inserted between PacI/AgeI
pCS2015	pCS1956 with CU LsIII ctrl trans inserted between PacI/AgeI
pCS2010	pCS1954 with CU LsIII trans inserted between PacI/AgeI
pCS2011	pCS1954 with CU LsIII ctrl trans inserted between PacI/AgeI
pCS2008	pCS1953 with CU LsIII trans inserted between PacI/AgeI
pCS2009	pCS1953 with CU LsIII ctrl trans inserted between PacI/AgeI
pCS2129	pcDNA5/FRT with DsRed-Express inserted between BamHI/AgeI
pCS2059	pCS2129 with tRNA 5' - CU LsIII trans - tRNA 3' inserted between BglII/MfeI
pCS2060	pCS2129 with tRNA 5' - CU LsIII ctrl trans - tRNA 3' inserted between BglII/MfeI
pCS2061	pCS2129 with tRNA 5' - HIV trans - tRNA 3' inserted between BglII/MfeI
pCS2062	pCS2129 with tRNA 5' - HIV ctrl trans - tRNA 3' inserted between BglII/MfeI
pCS1820	pCS1036 with sTRSVctrl with A spacers inserted between XhoI/ApaI
pCS1819	pCS1036 with sTRSV with A spacers inserted between XhoI/ApaI
pCS1816	pCS1036 with K with A spacers inserted between XhoI/ApaI
pCS1817	pCS1036 with W with A spacers inserted between XhoI/ApaI
pCS1818	pCS1036 with Y with A spacers inserted between XhoI/ApaI
pCS2080	pCS1036 with CU with B spacers inserted between XhoI/ApaI
pCS2081	pCS1036 with CU LsIII with B spacers inserted between XhoI/ApaI
pCS2083	pCS1036 with CU LsIII inversion with B spacers inserted between XhoI/ApaI
pCS1930	pCS1036 with CK LsI with A spacers inserted between XhoI/ApaI
pCS1931	pCS1036 with CK LsIII with A spacers inserted between XhoI/ApaI
pCS1932	pCS1036 with CK LsIV with A spacers inserted between XhoI/ApaI
pCS2085	pCS1036 with U LsIII with B spacers inserted between XhoI/ApaI
pCS2087	pCS1036 with HHe-PLMVd with B spacers inserted between XhoI/ApaI

pCS2088	pCS1036 with 3-way AA with B spacers inserted between XhoI/ApaI
pCS2089	pCS1036 with 3-way AAA with B spacers inserted between XhoI/ApaI
pCS2147	pCS1966 with d2EGFP inserted between HindIII/AvrII

Supplementary Table 2.3. Human cell lines with stably integrated constructs.

Parental line	Integrated plasmid construct
Flp-In HEK293	pCS1631
Flp-In HEK293	pCS1632
Flp-In HEK293	pCS1629
Flp-In HEK293	pCS1630
Flp-In HEK293	pCS1633
Flp-In HEK293	pCS1634
Flp-In HEK293	pCS1966
Flp-In HEK293	pCS2063

Chapter 3

Development of an RNA device framework that responds to proteins in human cells

Partially adapted from d'Espaux, L. D., Kennedy, A. B., Vowles, J. V., & Smolke, C. D. (2014). Development of protein-responsive ribozyme switches in eukaryotic cells. In preparation.

Abstract

Synthetic molecular devices for programmable gene regulation in human cells are useful tools for studying biological systems and for developing novel diagnostic and therapeutic platforms. Ribozyme switches are a class of gene-regulatory device that have been designed to exhibit programmable gene regulation activity in response to small molecule ligand inputs; however, response to protein ligands has not been demonstrated to date. We developed ribozyme switches that respond to the bacteriophage MS2 coat protein and demonstrated ligand-responsive modulation of gene expression in a human cell line. We investigated different strategies for device architecture and optimized the protein ligand to maximize sensitivity of the system, demonstrating up to 6.5-fold activation and up to 4.6-fold inhibition of gene expression from the ON and OFF switch platforms, respectively. We also explored the mechanism of action and ligand localization requirements of the ribozyme switch by localizing the protein ligand to different cellular compartments. We found that ligand localization to either the nucleus or the cytoplasm is sufficient for switching activity. Finally, we extended the platform to the design of a ribozyme switch responsive to the endogenous signaling protein β -catenin, demonstrating the ability of our platform to respond to an important disease marker.

Introduction

Synthetic ligand-responsive genetic regulators are important tools for controlling diverse biological systems spanning from engineered microorganisms to human patients. Many small-molecule-responsive RNA-based switches have been demonstrated in eukaryotic systems, but fewer have responded to protein inputs¹. As changes in protein expression determine cellular phenotype, molecular devices that directly detect and respond to intracellular concentrations of proteins are important engineering tools. Such genetically encoded devices have applications in the noninvasive detection and quantification of proteins in a complex cellular environment, as well as in targeting therapeutic activities to specific diseased cellular states.

Several examples of mammalian gene control platforms have been described that utilize different architectures and gene-regulatory mechanisms. One common strategy is the placement of an aptamer in the 5' untranslated region (UTR) of the gene of interest, such that ligand binding causes translational repression by preventing proper ribosome assembly for initiating translation²⁻⁷. In one example, binding of tryptophan RNA-binding attenuation protein (TRAP) to its 5' UTR binding site produced 180-fold translational repression². In another example, the archaeal ribosomal protein L7Ae was used to regulate two reporter genes simultaneously⁶. However, this approach is limited to OFF switches unless it is coupled with an additional genetic inverter component⁸, and can respond only to cytoplasmic protein ligands. Additionally, in mammalian cells 5' UTR secondary structure can nonspecifically interfere with translation, further limiting this regulatory strategy.

III-4

The incorporation of other classes of gene-regulatory elements can further expand the capabilities of these protein-responsive switches. For example, an L7Ae-responsive ON switch was created using a trans-acting regulator RNA that bound to the translation initiation site on the target messenger RNA (mRNA), thereby silencing gene expression⁴. Binding of L7Ae to its aptamer in the regulator RNA prevented binding to the mRNA, thereby derepressing expression of the target gene. As another example, an L7Ae-responsive ON switch based on an RNA interference (RNAi)-based silencing gene-regulatory element was described that incorporated an aptamer in the loop region of the short hairpin RNA (shRNA), such that ligand binding in the cytoplasm masked the Dicer recognition site and prevented processing, inhibiting the gene silencing effect observed with proper Dicer processing in the absence of ligand⁹. In a third example, ON and OFF switches modulating precursor mRNA (pre-mRNA) alternative splicing were used to control the inclusion of an exon containing a premature stop codon in response to endogenous protein disease markers in the nucleus¹⁰.

Protein-responsive translational regulators have also been used to regulate genetic circuits in mammalian cells. In one example, an L7Ae-responsive OFF switch was used to control L7Ae expression in a genetic feedback loop⁵. In another example, that same switch was used in conjunction with an shRNA-based ON switch to precisely regulate the relative expression levels of a proapoptotic and an antiapoptotic gene, thereby controlling cell fate in HeLa cells⁹. In addition, more complex genetic circuitry has been demonstrated by incorporating both protein-responsive translational regulators and small-molecule-responsive transcriptional regulators to engineer genetic systems that perform

digital logic computations and fundamental arithmetic operations on prescribed molecular inputs³.

One main advantage of post-transcriptional regulators such as those described above relative to transcriptional regulators is that they exhibit a faster change in gene expression in response to changes in the level of protein ligand. When translation of the transcript is hindered or the transcript is degraded, synthesis of the encoded protein ceases, while the product of a transcriptionally silenced gene continues to be translated from existing mRNA. However, generally the components in these regulatory devices are capable of either ON or OFF switching, but not both. Rarely can the components be coupled without disrupting their individual functions, and the devices are generally not portable between microbes and higher eukaryotes. A single gene-regulatory device platform capable of overcoming these limitations would present a more flexible and streamlined design process for devices tailored to different systems.

Ligand-responsive ribozyme switches have been used to regulate gene expression in yeast¹¹⁻¹³ and mammalian cells¹⁴⁻¹⁶. Ribozyme switches are incorporated into the UTR of the target gene, where ligand-regulated cleavage of the transcript leads to silencing of gene expression¹¹. This platform possesses a number of important advantages that many of the protein-responsive devices demonstrated to date in mammalian cells lack. First, ribozyme switches can be programmed to turn gene expression on or off in response to almost any ligand for which there exists an RNA aptamer. Second, switch activity can be tuned through modifications to the sequence of the aptamer, ribozyme, and transmitter components. Third, switch components can be modularly coupled without disrupting their activities. Finally, ribozyme switches are

highly portably between different organisms because their mechanism of action is independent of cell-specific machinery^{11,17}. However, one current limitation of ribozyme switches is that the platform has only been demonstrated to respond to small molecule ligands. It is also unknown where in the cell the ribozyme cleavage event occurs. A protein-responsive ribozyme switch platform would leverage the important design advantages of the existing small-molecule-responsive platform while expanding its capability to process a new and important class of inputs.

We describe the development of a protein-responsive ribozyme switch platform for regulating gene expression in mammalian cells. We investigated a variety of device architectures, expanding beyond the design of the small-molecule responsive ribozyme switches, using an aptamer for the bacteriophage MS2 coat protein and controlling the expression of a fluorescent reporter gene. We developed and optimized a genetic expression system for quantitative characterization of device activity in human cells. The most highly active switch designs exhibited up to 6.5-fold activation and up to 4.6-fold inhibition of gene expression from the ON and OFF switch architectures, respectively. Experiments examining the impact of ligand localization on device activity indicated that the ribozyme switch platform is uniquely flexible in responding to ligands in either the nucleus or the cytoplasm. Finally, we describe attempts to develop ribozyme switches responsive to other proteins and demonstrate a device responsive to the endogenous signaling protein β -catenin.

Results

Design of protein-responsive ribozyme switch platforms

While previous work in the Smolke laboratory has demonstrated ligand-responsive ribozyme switches, these studies have been limited to small molecule ligands¹¹. In developing protein-responsive ribozyme switch platforms we focused our initial designs on integration strategies for an aptamer to the bacteriophage MS2 coat protein (sensor) and the satellite RNA of tobacco ringspot virus (sTRSV) hammerhead ribozyme (actuator). The MS2 protein and its natural stem-loop aptamer were selected as the initial ligand-aptamer pair for this study as they have been extensively used in cellular systems^{2,3,6,18-22} and as MS2 is a heterologous protein its levels can be readily controlled. Using this sensor and actuator we explored different design strategies for coupling these components in a way that would enable the binding of the protein ligand to affect the activity of the hammerhead ribozyme.

The simplest design we explored was the direct-coupled architecture, in which the MS2 aptamer is directly coupled to either loop I or loop II of the hammerhead ribozyme (Figure 3.1, MS2-A), without any separate transmitter component linking the sensor and the actuator as incorporated in other small-molecule-responsive ribozyme switch platforms¹¹. We designed two different aptamer integration points in each of the loops, and we varied the loop I sequence. The hypothesis underlying the direct-coupled architecture is that the relatively large size of the protein ligand MS2 would disrupt the ability of the ribozyme to fold into the catalytically-active three-dimensional structure, sterically interfering with the tertiary interactions between loops I and II that have been shown to be necessary for cleavage activity at physiological Mg^{2+} concentrations²³. The

III-8

direct-coupled architecture obviates the need to design the ribozyme switch to adopt distinct cleavage-active and cleavage-inactive conformations, and thus is a simpler design than strategies that incorporate directed secondary structure rearrangements. However, it is possible that this design may be more dependent on the size of the protein ligand and the orientation by which the protein ligand binds to its aptamer.

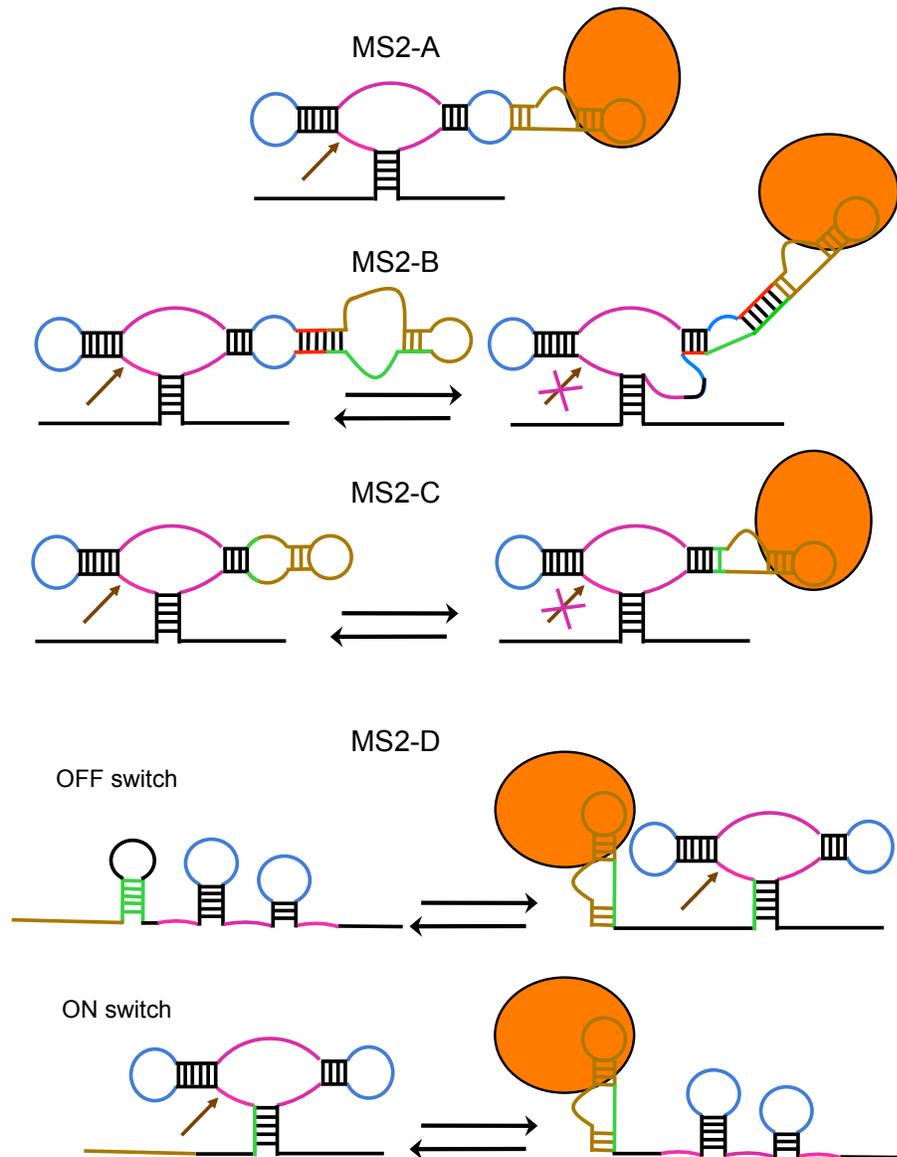


Figure 3.1. MS2-responsive ribozyme switch designs. MS2-A: The MS2 aptamer is coupled directly to the loop. MS2-B: The MS2 aptamer is coupled through a transmitter affecting the secondary structure of stem II and the catalytic core. MS2-C: The MS2 aptamer is coupled through a transmitter affecting the secondary structure of the loop. MS2-D: The MS2 aptamer is placed directly upstream of the ribozyme and a transmitter affecting the formation of stem III. The catalytic core is shown in magenta, loops and bulges are shown in blue, stems are shown in black, the transmitter is shown in green and red, the aptamer is shown in brown, and the MS2 ligand is shown in orange. The cleavage site is indicated with an arrow. See Supplementary Table 3.1 for sequences.

We designed three additional ribozyme switch platforms that incorporated directed secondary structure rearrangements into distinct cleavage-active and cleavage inactive conformations. The first set of ribozyme switches contains a transmitter designed to alter the secondary structure of one of the stems and the catalytic core (Figure 3.1, MS2-B). These ribozyme switches employ the same design strategy as previously developed small-molecule-responsive devices¹¹, with sequence variations in loop I, the transmitter, and the aptamer. The next set of designs contains a loop-transmitter that alters the secondary structure of the loop to which the aptamer is attached (Figure 3.1, MS2-C), which is expected to disrupt important tertiary interactions with the other loop. We designed an OFF switch, MS2-C1, in which the aptamer and loop-transmitter are integrated into loop I, and three ON switches with varied loop I sequences in which the aptamer and loop-transmitter are integrated into loop II. In the final set of designs, the aptamer is not incorporated into the ribozyme, but is instead placed immediately

upstream in a sequential fashion (Figure 3.1, MS2-D). Alternate hybridization between the aptamer stem and stem III of the ribozyme prevents the aptamer and ribozyme from folding simultaneously in ON switches, while OFF switches contain a competing hairpin that prevents the folding of the aptamer and ribozyme, unless ligand binding stabilizes the aptamer and prevents the competing hairpin from folding. These six designs varied in the length and sequence identity of the transmitter component.

Initial characterization of protein-responsive ribozyme switches in a human cell line

The MS2-responsive ribozyme switch designs were tested for gene-regulatory and ligand-responsive activity in a human cell line. The initial characterization construct was based on the ribozyme switch regulating a green fluorescent protein (GFP) reporter protein and measuring fluorescence under different MS2 levels through a flow cytometry assay. Briefly, the ribozyme switches were located in the 3' UTR of destabilized enhanced GFP (d2EGFP)²⁴ expressed from a cytomegalovirus (CMV) promoter (Figure 3.2). A doxycycline-inducible expression cassette for MS2 was located on the same plasmid, in which an MS2-DsRed fusion protein was under the control of a CMV promoter with two tetracycline operator (TetO) sites located downstream of the promoter (CMV-TetO₂). The plasmid constructs were transiently transfected into a Flp-In T-REx human embryonic kidney 293 (HEK293) cell line, which stably expresses the tetracycline repressor (TetR). Thus, in the absence of doxycycline transcription of the protein ligand is inhibited by TetR, and transcription is activated by the addition of doxycycline which inhibits TetR binding to the operator sites.

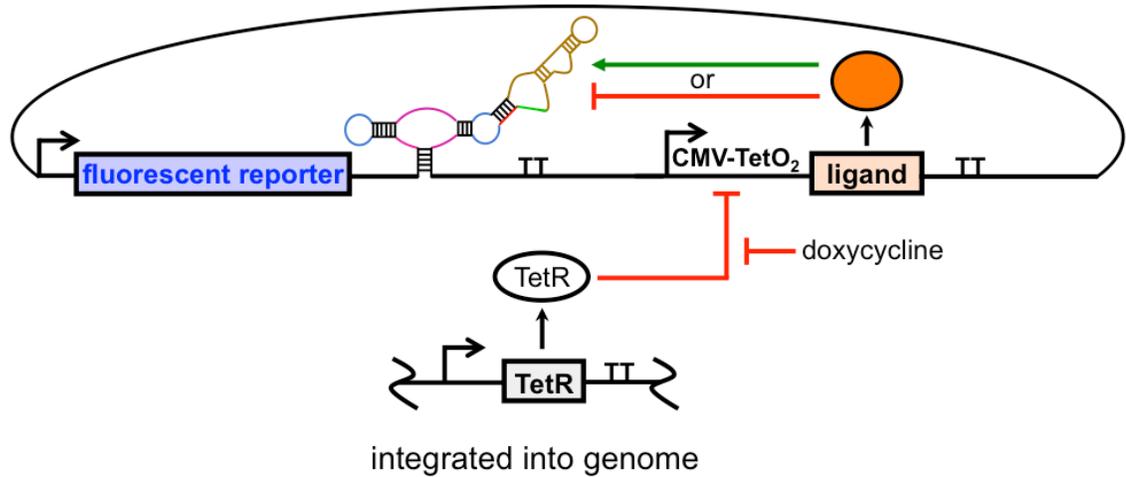


Figure 3.2. Protein-responsive ribozyme switch characterization system. A fluorescent reporter with ribozyme switch and a protein ligand are encoded on a plasmid. A tetracycline-responsive CMV-TetO₂ promoter controls expression of the protein ligand and the fluorescent reporter protein is expressed from a constitutive promoter. The plasmid is transiently transfected into a Flp-In T-REx HEK293 cell line, which expresses the TetR repressor. Addition of doxycycline derepresses the CMV-TetO₂ promoter, turning on expression of the ligand, which regulates the activity of ON and OFF switches. The plasmid (with both fluorescent reporter and ligand genes) can be used in transient transfection assays or stably integrated into the genome for stable expression assays.

We tested four of the direct-coupled designs (MS2-A1, MS2-A2, MS2-A5, and MS2-A6) in the described *in vivo* characterization assay. We selected designs that exhibited the highest *in vitro* cleavage rates in a surface plasmon resonance (SPR) assay (Andrew Kennedy, unpublished results). We found that these ribozyme switch designs were able to downregulate GFP gene expression to varying degrees, but that cleavage activity was not attenuated by MS2 (Figure 3.3). Subsequent *in vitro* switching assays

(Andrew Kennedy, unpublished results) confirmed that although these designs were capable of both cleavage and binding to MS2, ligand binding did not diminish cleavage activity. These results suggest that although MS2 is approximately the same size as the ribozyme switch, binding alone is not sufficient to disrupt tertiary interactions between the loops and thus cleavage activity.

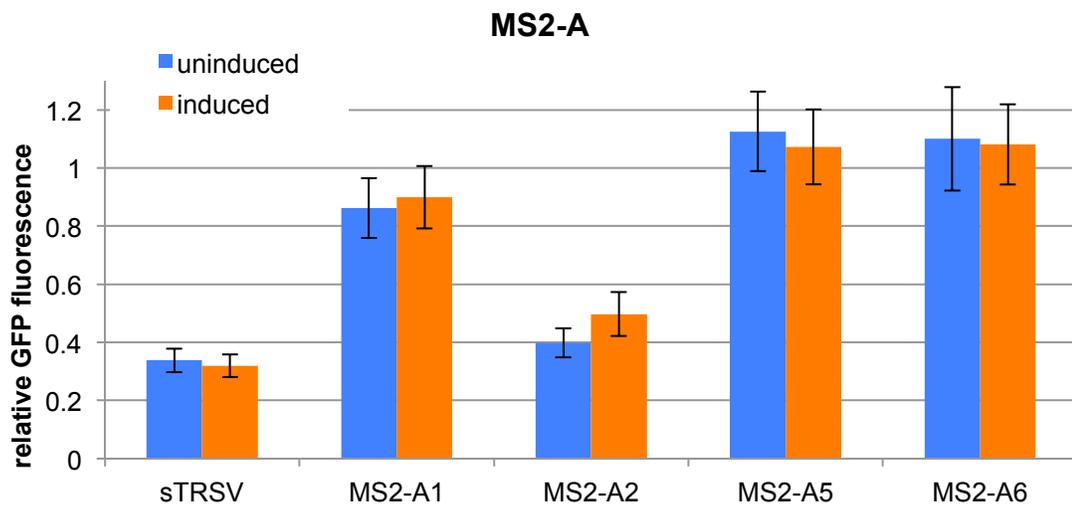


Figure 3.3. Activity of MS2-A designs. Relative GFP fluorescence levels are reported for transiently transfected constructs encoding ribozyme switch sequences. Reported values are geometric mean \pm s.d. from biological duplicates and normalized to the absence of ribozyme.

All other ribozyme switch platforms were tested using a slightly modified *in vivo* characterization construct. Specifically, the d2EGFP fluorescent reporter gene was replaced with blue fluorescent protein (BFP), the CMV promoter driving the expression of the reporter gene was replaced with the elongation factor 1 α (EF1 α) promoter, and the MS2-DsRed fusion ligand was replaced with MS2 (Figure 3.2). Flow cytometry assays

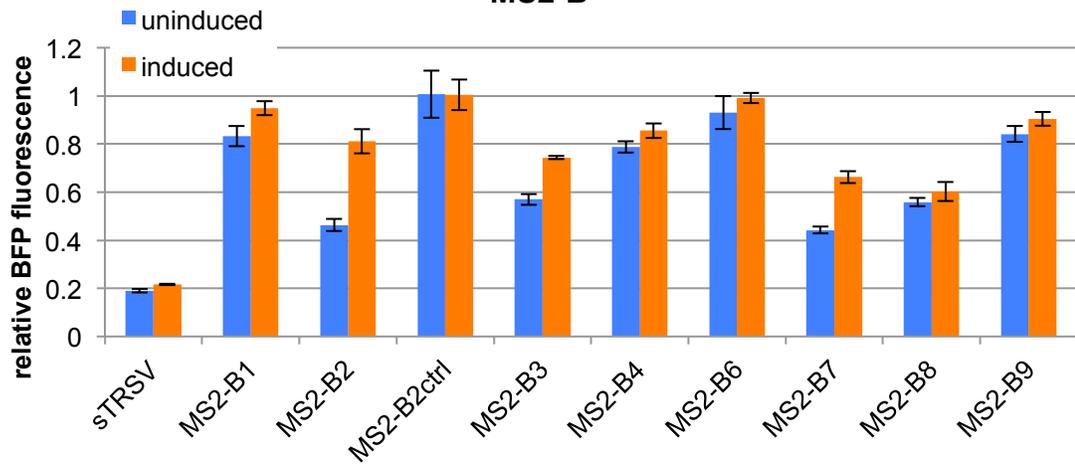
were performed on a Flp-In T-REx HEK293 cell line transiently transfected with the ribozyme switch constructs. The transmitter designs (MS2-B) showed low to moderate amounts of gene knockdown activity, with three designs, MS2-B2, MS2-B3, and MS2-B7, exhibiting 1.8-, 1.3-, and 1.5-fold, respectively, increases in gene expression in response to doxycycline-induced MS2 (Figure 3.4, MS2-B). MS2-B2 showed the highest switching activity of all designs tested, although its high basal expression level may limit its usefulness in future applications.

One of the loop-transmitter designs (MS2-C) showed low gene knockdown and no switching activity (MS2-C2), while the other three designs responded to MS2 (Figure 3.4, MS2-C). MS2-C3 exhibited a high level of gene knockdown activity, with gene expression almost as low as sTRSV, and the second highest switching activity (1.6-fold increase). MS2-C1 was the only functional OFF switch tested, exhibiting 1.4-fold reduction in gene expression in response to MS2.

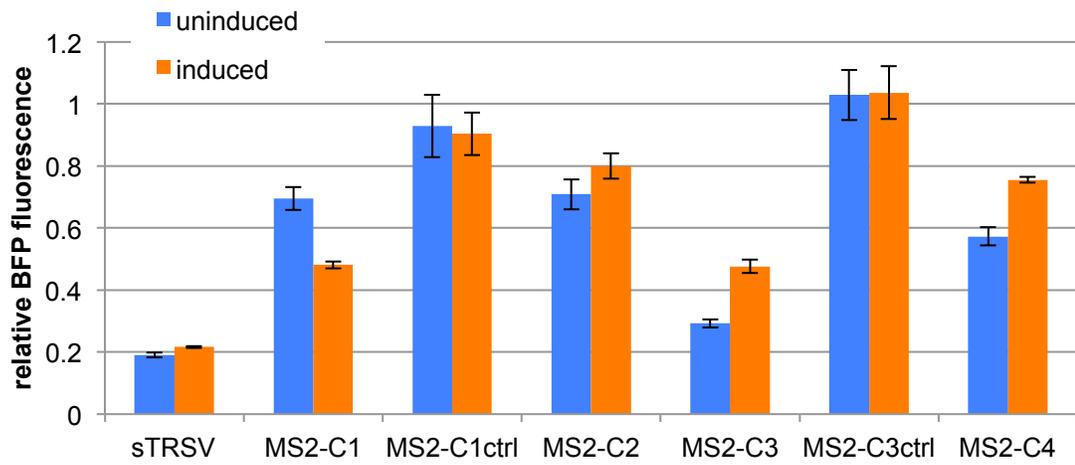
The sequential designs (MS2-D) showed the highest levels of gene knockdown activity, with three designs, MS2-D3, MS2-D5, and MS2-D6, exhibiting gene expression as low as sTRSV (Figure 3.4, MS2-D). MS2-D5 showed the highest switching activity (1.5-fold) of the sequential designs.

III-14

MS2-B



MS2-C



MS2-D

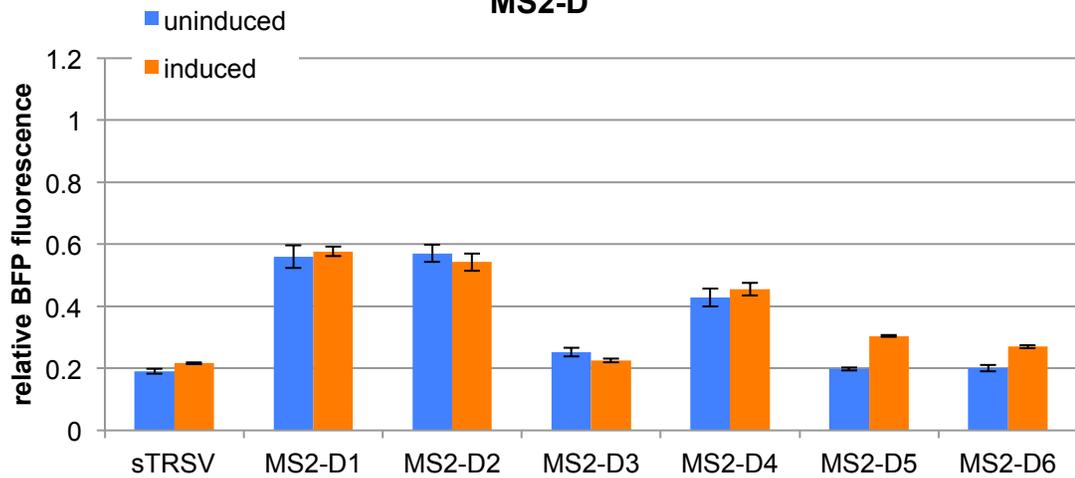


Figure 3.4. Activity of MS2-B, MS2-C, and MS2-D designs. Relative BFP fluorescence levels are reported for transiently transfected constructs encoding ribozyme switch sequences. Reported values are geometric mean \pm s.d. from biological duplicates and normalized to the non-cleaving sTRSVctrl.

Non-cleaving control versions of three ribozyme switches (MS2-B2ctrl, MS2-C1ctrl, and MS2-C3ctrl) were constructed and tested under identical assay conditions. The control designs did not exhibit gene knockdown activity or responsiveness to MS2 (Figure 3.4). These results indicate that gene knockdown is due to ribozyme cleavage rather than to an effect of device secondary structure alone, and that ligand responsiveness is due to modulation of cleavage rather than an effect of ligand binding alone. Eight designs (MS2-B1, MS2-B2, MS2-B3, MS2-B7, MS2-C1, MS2-C3, MS2-C4, and MS2-D5) were subjected to further testing.

Development of an improved genetic system for quantitative characterization of ribozyme switch activity *in vivo*

In order to more accurately measure the gene-regulatory activity of the protein-responsive ribozyme switches in mammalian cells, we developed and optimized an improved *in vivo* characterization system. As described above, the initial characterization system for the ribozyme switches (Figure 3.2) measured activities through transient transfection assays in order to quickly screen device designs for activity. However, the ease of this screening method is accompanied by a high degree of variability. For example, in a transient assay different cells in the transfected population receive different

amounts of plasmid, such that cells in the transfected population can exhibit a range of gene expression over three orders of magnitude, and a large portion of cells receives no plasmid at all (Figure 3.5). To measure ribozyme switch activity, analysis must be performed on transfected cells only, which was accomplished by gating for cells that expressed a fluorescent protein as a transfection marker. One method that we used to determine the transfected population was to cotransfect a plasmid encoding the expression of a second fluorescent protein with the plasmid encoding a ribozyme switch and its reporter gene. Control experiments in our laboratory have shown that in such cotransfection experiments nearly all transfected cells contain both plasmids (Kathy Wei and Joy Xiang, unpublished results). An alternative method we used to determine the transfected population was to gate transfected cells based on the expression of the reporter gene regulated by the ribozyme switches. While cell populations exhibiting a high level of gene expression are fully distinct from populations of untransfected cells, this is not so for cells expressing devices with high gene silencing activity, with some transfected cells exhibiting similar fluorescence levels as untransfected cells (Figure 3.5). Such devices therefore exhibit an artificially high level of gene expression after gating by this method, as the transfected cells with the lowest gene expression levels have been removed from the analysis. While this effect may result in elevated expression levels for these switches, it does not obscure our ability to observe the switching activity associated with these switches, even for devices with the lowest levels of gene expression (Figure 3.4, MS2-D5 and MS2-D6).

To avoid the variability and efficiency issues inherent in transient transfections, we examined the characterization of constructs that had been stably integrated into the

genome of the cell line as an alternative approach. We used the Flp-In integration system with an HEK293 cell line containing a Flp Recombination Target (FRT) site in its genome. This system allows an expression vector to be integrated into the genome via Flp recombinase-mediated recombination at the FRT site. The resulting isogenic stable cell line presents a homogenous level of gene expression across all cells in the population, although this level is lower than that produced by the same construct when transiently transfected (Figure 3.5). As cells with the desired integration are selected by culturing the transfected cells in selective media, the integration process requires 2–3 weeks to generate isogenic stable cell lines. To ensure maximum flexibility in ribozyme switch characterization, the characterization plasmids were constructed on a backbone that was compatible with both transient transfection and stable integration procedures (Figure 3.2).

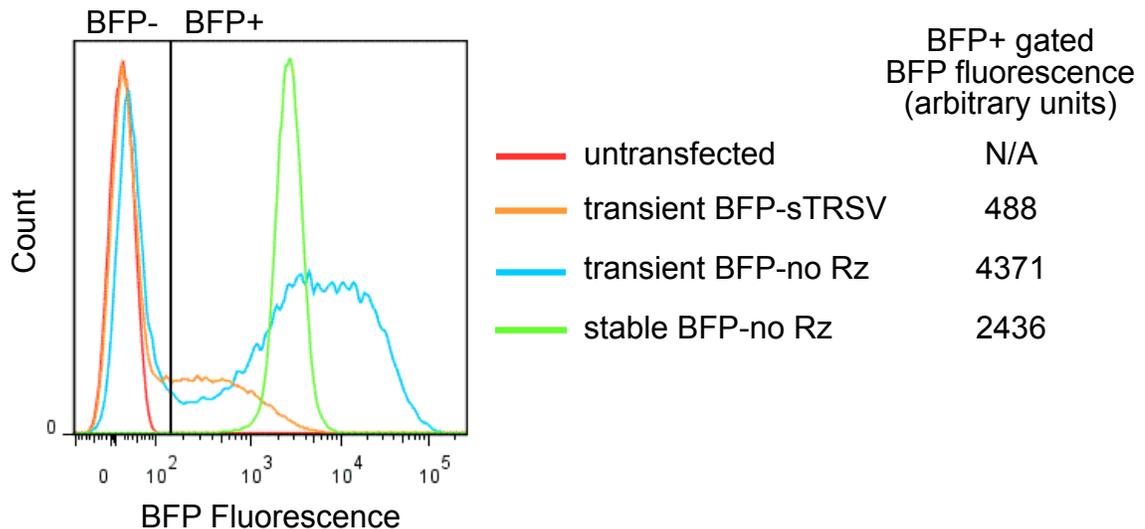


Figure 3.5. Histograms of transiently transfected and stably integrated fluorescent constructs. Transient transfection is not highly efficient and results in a large population

of untransfected cells (red). Gating to remove those cells from analysis also removes a portion of a population of cells expressing a device with a high level of gene silencing (orange). Stable integration of a construct yields a highly homogenous population (green) with lower mean fluorescence than cells transiently transfected with that construct (blue).

As described above, our initial characterization construct utilized a GFP reporter expressed from a CMV promoter to measure the gene-regulatory activity of the ribozyme switches. These constructs encoded a protein ligand expression cassette in which an MS2-DsRed fusion was expressed from a doxycycline-inducible CMV-TetO₂ promoter. We observed that when these constructs were stably integrated into Flp-In T-REx HEK293 cells, the resulting stable cell lines produced no detectable fluorescent signal from the MS2-DsRed fusion in response to doxycycline. In contrast, when using the same constructs in transient transfection assays, doxycycline did elicit a detectable fluorescent signal, albeit low, from the MS2-DsRed fusion. Our control experiments indicated that the expression from the stably integrated constructs were generally lower than that from the same constructs in a transient transfection assay (Figure 3.5). Thus, we concluded that under stable expression conditions the levels of the MS2-DsRed ligand were reduced below the detection threshold of the flow cytometry assay.

To improve characterization of the protein responsiveness of the ribozyme switches in the stable expression assay, we modified the protocol to increase protein ligand levels by transiently transfecting a plasmid encoding the expression of MS2 or the MS2-DsRed fusion into the stable cell lines prior to analysis. However, in the course of

optimizing the protocol for this assay, we noticed two unexpected effects that confounded analysis of the gene-regulatory activity and ligand-responsiveness of the ribozyme switches. First, transient transfection of d2EGFP-expressing cell lines resulted in a population of cells with lower GFP fluorescence levels than untransfected cells, but not as low as parental cells expressing no fluorescent proteins (Figure 3.6A). The appearance of this cell population was observed for transfection of plasmids encoding expression of fluorescent proteins, non-fluorescent proteins, and no proteins (i.e., a plasmid with no mammalian promoters). Transfection of a plasmid encoding a different fluorescent reporter protein, BFP, under the control of a different promoter, EF1 α , also led to this effect, ruling out promoter competition as a cause of the knockdown. Importantly, this effect was not observed when an “empty” transient transfection was performed, containing transfection reagent but no plasmid.

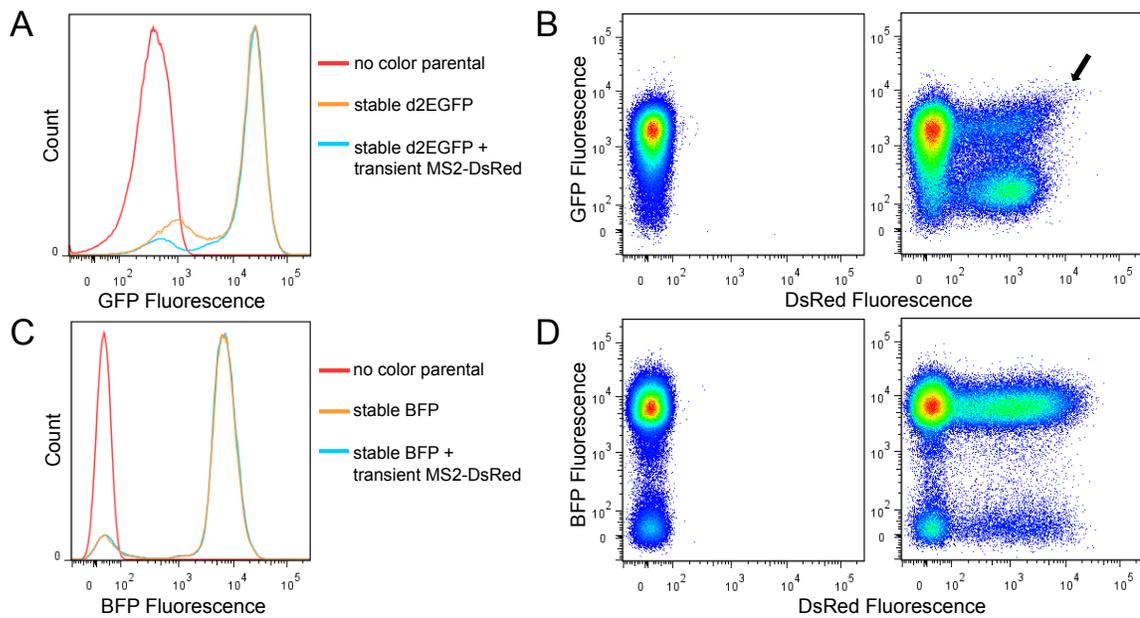


Figure 3.6. Effects of transient transfection of stable cell lines expressing a fluorescent reporter protein. **(A)** Histograms of a d2EGFP stable cell line (blue), that stable line transfected with a plasmid encoding MS2-DsRed (orange), and the no color parental cell line (red). **(B)** Two color scatter plots of a d2EGFP stable cell line, showing the small population of cells (arrow) transfected with a plasmid encoding MS2-DsRed that exhibit higher GFP fluorescence than the stable line itself. **(C)** Histograms of a BFP stable cell line (blue), that stable line transfected with a plasmid encoding MS2-DsRed (orange), and the no color parental cell line (red). **(D)** Two color scatter plots of a BFP stable cell line transfected with a plasmid encoding MS2-DsRed.

The second confounding effect resulting from transient transfection of plasmids encoding MS2-DsRed into the cell lines stably expressing d2EGFP was an increase in GFP fluorescence in cells with the highest DsRed fluorescence to levels higher than the untransfected cell population (Figure 3.6B). Such an effect can be expected from spillover between fluorescence channels due to overlap of the fluorescent protein emission spectra; however, compensation is generally performed to correct for such effects. In this case, the effect remained after proper compensation, and even after over-compensation, suggesting that it is not due to spillover between fluorescence channels. This effect is especially problematic as it is similar to the changes we would expect to observe in GFP fluorescence as a result of the gene-regulatory activity of an MS2-responsive ribozyme switch, with ligand binding to the ribozyme switch preventing cleavage and thereby increasing gene expression. However, the effect was also observed with plasmids encoding DsRed without the MS2 ligand, and in stable lines that did not

contain a ribozyme switch. We therefore concluded that the increase in GFP fluorescence was not a result of ribozyme switch activity.

To address both of these problems, we redesigned our characterization system such that the ribozyme switches were placed in the 3' UTR of BFP, which was under the control of an EF1 α promoter. Stable cell lines expressing BFP did not exhibit nonspecific knockdown or fluorescence increases as a result of transient transfection of plasmids, including plasmids encoding the MS2-DsRed fusion (Figure 3.6C and D). This redesigned and optimized characterization system was used for all subsequent ribozyme switch characterization assays, including investigations of the transmitter designs (MS2-B), the loop-transmitter designs (MS2-C), and the sequential designs (MS2-D).

MS2 variants result in optimization of switch sensitivity to ligand

In our initial screening we observed a moderate amount of switching activity from a subset of our MS2-responsive ribozyme switch designs. We next examined whether we could improve ligand sensitivity, and thus switching activity, of these designs by optimizing the ligand itself. MS2 binds to its aptamer in the dimerized form²⁵. However, once bound the wild-type protein will multimerize to form a capsid²⁶. We hypothesized that multimerization of the ligand could negatively impact switching activity. Thus, we examined two alternative versions of the MS2 ligand: (i) a mutant form of MS2 (MS2mut) containing two amino acid substitutions (V75E and A81G) that is deficient in capsid formation but retains the RNA binding affinity of the wild-type protein²⁷ and (ii) a fused dimer of the MS2 mutant (2MS2mut). The dimer forms in a head-to-tail orientation, and the fused dimer joins the N- and C-termini together with the deletion of 3

amino acids at the junction. A similar fused dimer without the V75E/A81G substitutions has been shown to retain the same RNA binding affinity as wild-type²⁸.

We assayed a subset of ribozyme switch designs for sensitivity to these MS2 variants using the improved characterization system (Figure 3.2). We tested the subset of designs that responded to MS2 in our initial screening experiments, as well as MS2-B1 because it showed substantial switching activity in yeast assays (Leopold d’Espaux, unpublished results). We performed the assay using transient expression, similar to our initial experiments with the MS2-B, MS2-C, and MS2-D designs, with MS2mut or 2MS2mut replacing MS2 on the plasmid. Our results indicate that although device basal levels remained mostly unaffected by the change in ligand (as expected), all switches tested were equally or more responsive to MS2mut than MS2 (Figure 3.7A). In addition, most of the switch designs exhibited even greater sensitivity to the 2MS2mut ligand, including MS2-B1, which showed no response to MS2 or MS2mut under identical assay conditions. These results indicate that preventing ligand multimerization is beneficial for switch sensitivity. The increase in sensitivity to 2MS2mut is likely due to increased effective ligand concentration. As the MS2 monomer must dimerize in order to bind to the aptamer²⁵, expressing the protein as a fused dimer roughly doubles the effective concentration of the ligand.

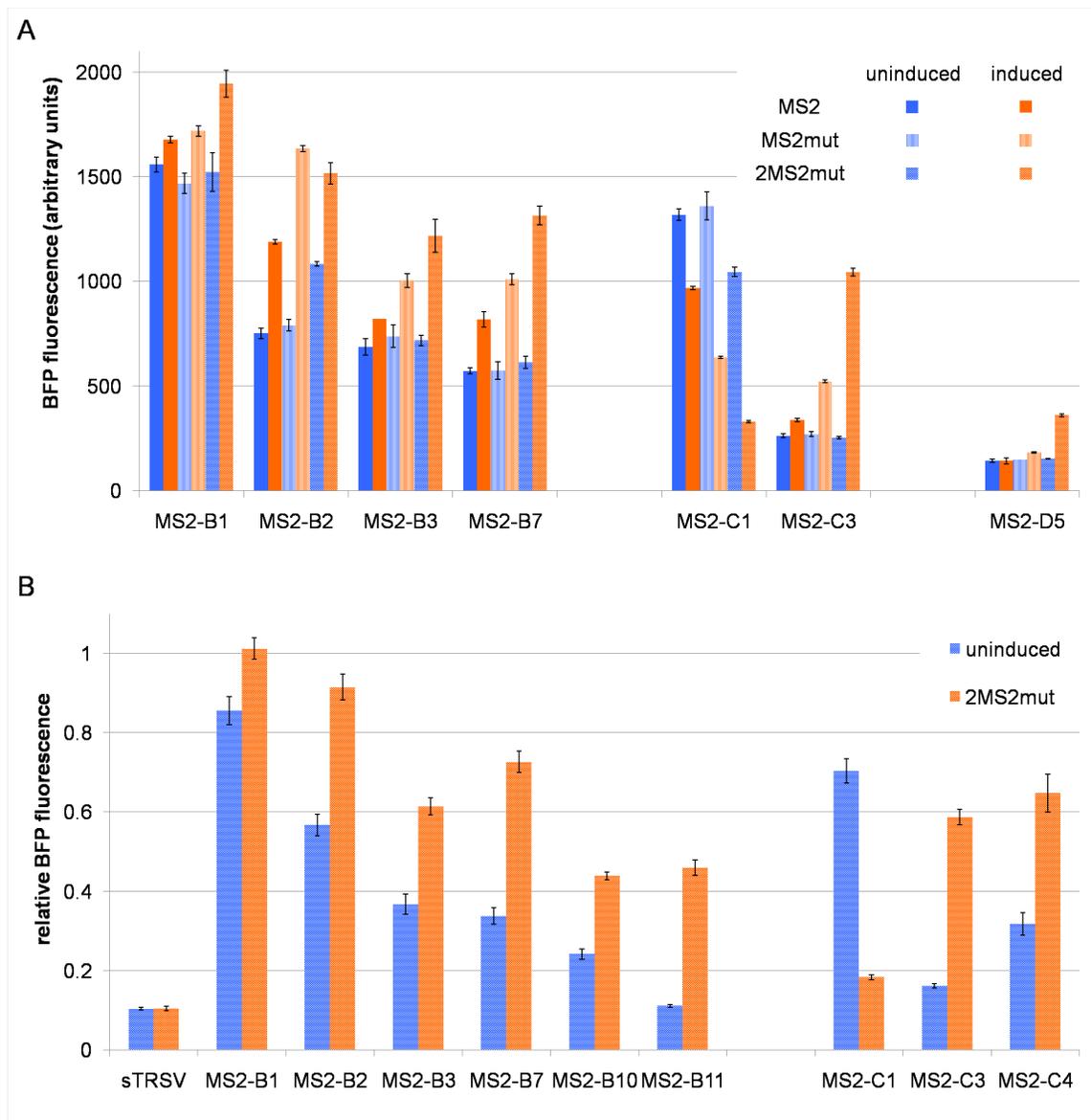


Figure 3.7. Activity of ribozyme switches with optimized ligands. **(A)** Switch response to wild-type MS2, MS2 V75E/A81G (MS2mut), and the fused dimer of MS2mut (2MS2mut). BFP fluorescence levels are reported for transiently transfected constructs encoding ribozyme switch sequences. Reported values are geometric mean \pm s.d. from biological duplicates. **(B)** Switch response to 2MS2mut, including two new switch designs. Relative BFP fluorescence levels are reported for transiently transfected constructs encoding ribozyme switch sequences cotransfected with a transfection control

plasmid. Reported values are geometric mean \pm s.d. from biological duplicates and normalized to the non-cleaving sTRSVctrl.

We used the results from this assay and our optimized ligand to further explore improvements to the ribozyme switch designs. MS2-B2 differs from MS2-B1 in the sequence of loop I, and we hypothesized that this difference may be the cause of the increased gene knockdown and switching activities observed for MS2-B2. We modified the designs of MS2-B3 and MS2-B7 by changing their loop I sequences to that of MS2-B2, generating MS2-B10 and MS2-B11, respectively. We assayed these new designs with our optimized ligand 2MS2mut (Figure 3.7B). MS2-B11 exhibited the greatest response to ligand (4.1-fold), while the OFF switch MS2-C1 exhibited 3.8-fold switching.

Protein ligand localization allows probing of ribozyme switch mechanism of action

We next investigated how the subcellular localization of ligand affected ribozyme switch activity. For ribozyme switches responsive to small molecules previously developed in the Smolke laboratory¹¹ it is expected that the ligand will freely diffuse throughout the cell, available to bind to its aptamer over the entire lifetime of the mRNA, from transcription in the nucleus to translation in the cytoplasm. Proteins, however, are commonly localized to specific subcellular locations. Localization of the protein ligand enables the investigation of where in the cell the ribozyme cleaves, and consequently when it cleaves relative to mRNA nuclear export and translation. In the case of an ON switch, where ligand binding is required to prevent ribozyme cleavage, the ribozyme

switch may be unresponsive if the ligand is localized to one cellular compartment. If the protein is localized to the cytoplasm the ribozyme may cleave in the nucleus during or immediately following transcription, and if the protein is localized to the nucleus the ribozyme may cleave after export to the cytoplasm. We attempted to elucidate the ligand localization requirements of our MS2-responsive switches in order to discover where in the cell the ribozyme cleaves, which in turn will inform the choice of new protein ligands for ribozyme switches in the future.

For all experiments described above, we expected the ligand to be found in both the nucleus and the cytoplasm, as MS2 (14 kDa), MS2mut (14 kDa), and 2MS2mut (28 kDa) are small enough to passively diffuse through the nuclear pore without the aid of any nuclear transport machinery²⁹. We attempted to control protein localization by creating 2MS2mut constructs with either an N-terminal nuclear localization sequence derived from Simian virus 40 (SV40)³⁰ (NLS-2MS2mut) or a C-terminal nuclear export sequence derived from protein kinase A inhibitor α (PKI α)³¹ (2MS2mut-NES). We transiently transfected Flp-In T-REx HEK293 cells with plasmids encoding each of these three 2MS2mut localization variants, induced protein ligand expression from the CMV-TetO₂ promoter using doxycycline, and harvested the total protein in nuclear and cytoplasmic extractions. Immunoblotting of these extracts with an antibody specific for MS2 revealed that 80% of an average cell's 2MS2mut is found in the cytoplasm, despite its small size (Figure 3.8A). The NES tag localized 90% of an average cell's 2MS2mut to the cytoplasm, which is comparable to the distribution of the cytoplasmic control protein. Compared to the untagged version of 2MS2mut, the distribution of NLS-tagged protein was shifted towards the nucleus, but a significant amount remained in the

cytoplasm, such that the protein was present at approximately the same concentration in both compartments. Immunofluorescence microscopy of stable cell lines expressing the three localization variants of 2MS2mut using the same anti-MS2 antibody validated the subcellular distribution determined by immunoblotting (Figure 3.8B).

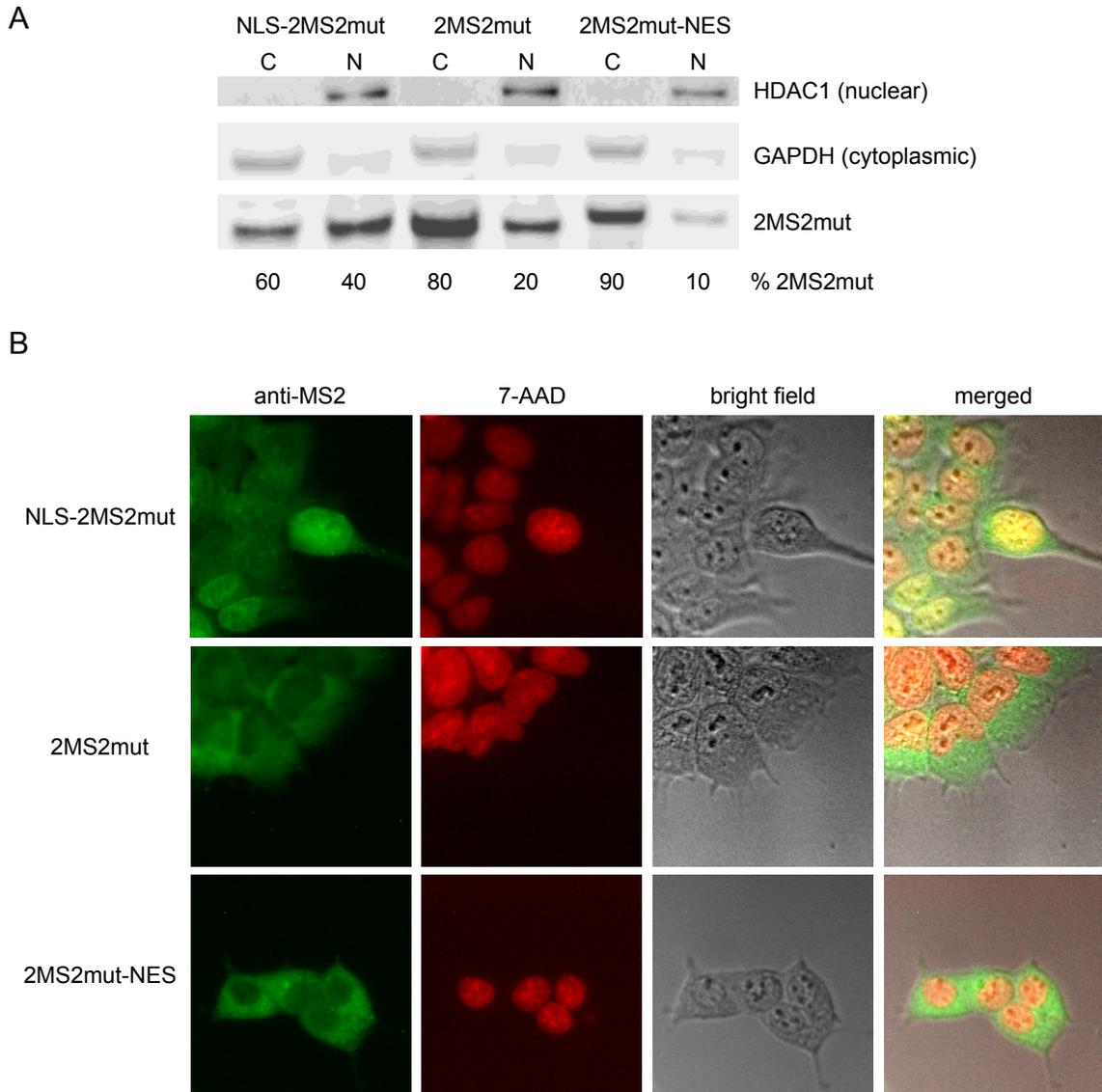


Figure 3.8. Subcellular localization of ligands with and without localization signals. (A) Immunoblot showing subcellular compartment distribution of 2MS2mut, NLS-2MS2mut,

and 2MS2mut-NES, along with nuclear and cytoplasmic controls. Percentages are calculated by normalizing quantified band intensity to the number of cells harvested. C, cytoplasm; N, nucleus; HDAC1, histone deacetylase 1; GAPDH, glyceraldehyde 3-phosphate dehydrogenase. **(B)** Immunofluorescence microscopy of 2MS2mut, NLS-2MS2mut, and 2MS2mut-NES using the same anti-MS2 antibody as in **A**. Green, anti-MS2 antibody and fluorescent secondary antibody; red, 7-aminoactinomycin D (7-AAD) nuclear stain.

We tested the localization-tagged versions of 2MS2mut with a subset of ribozyme switches in both transient and stable expression assays (Figure 3.9). While the switches responded similarly to 2MS2mut and 2MS2mut-NES, we observed that NLS-2MS2mut produced a slightly lower level of switching activity. As we have repeatedly observed by flow cytometry and immunoblotting that proteins with NLS tags are present at lower levels than untagged versions, we inferred that the lower switching activity was likely due to lower protein levels and not an effect of the altered subcellular distribution of the protein. Importantly, because the switches respond to 2MS2mut-NES, we concluded that cytoplasmic localization of the protein ligand is sufficient for switching activity, yielding as much as 5.3-fold ON switching with MS2-B11 and 4.3-fold OFF switching with MS2-C1 in a stable expression assay.

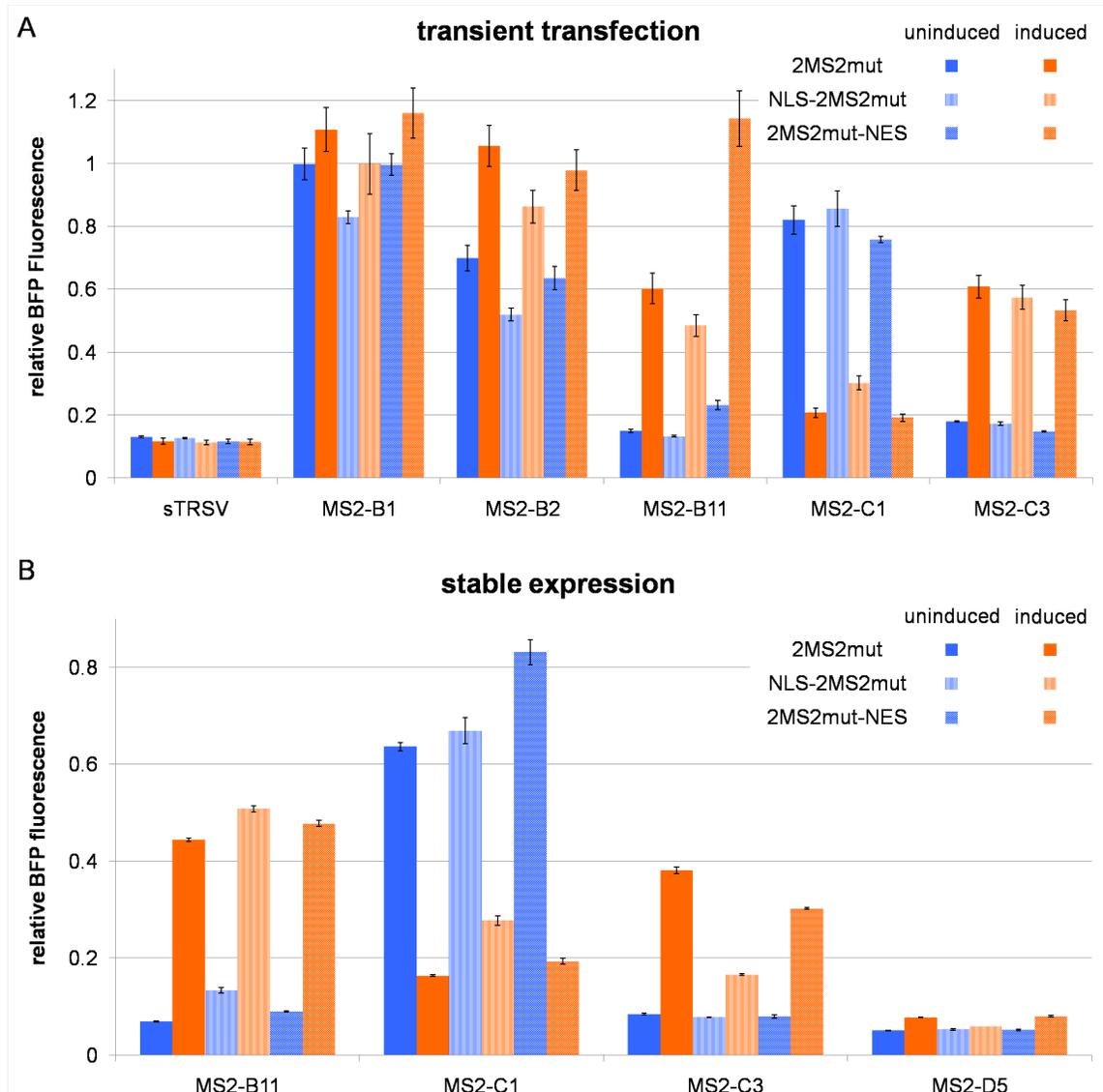


Figure 3.9. Activity of ribozyme switches with ligands with and without localization signals: 2MS2mut, NLS-2MS2mut, and 2MS2mut-NES. **(A)** Relative BFP fluorescence levels are reported for transiently transfected constructs encoding ribozyme switch sequences cotransfected with a transfection control plasmid. Reported values are geometric mean \pm s.d. from biological duplicates and normalized to the non-cleaving sTRSVctrl. **(B)** Relative BFP fluorescence levels are reported for stably integrated constructs encoding ribozyme switch sequences. Reported values are geometric mean \pm

s.d. from biological duplicates or triplicates and normalized to the non-cleaving sTRSVctrl.

As our NLS-2MS2mut protein failed to fully localize to the nucleus, we created new constructs to achieve the desired localization. We hypothesized that although NLS-2MS2mut was being actively transported into the nucleus, its small size allowed it to passively diffuse out of the nucleus and accumulate in the cytoplasm to a significant level despite the NLS. We created 2MS2mut-DsRed fusion proteins with either an N-terminal NLS (NLS-2MS2mut-DsRed) or a C-terminal NES (2MS2mut-DsRed-NES), as well as 2MS2mut-DsRed without a localization sequence. These larger fusion proteins were expected to passively diffuse through the nuclear pore to a much lower extent than 2MS2mut without DsRed, and the fluorescent tag allowed for direct MS2 detection without immunostaining. We transiently transfected Flp-In T-REx HEK293 cells with plasmids encoding each of these three protein variants, induced MS2 expression from the CMV-TetO₂ promoter using doxycycline, and imaged the cells using confocal fluorescence microscopy (Figure 3.10). All three variants of the protein localized to the expected cellular locations. The presence in these constructs of BFP, which does not contain a localization signal and was found in both the nucleus and the cytoplasm, allowed for direct comparison between this protein and DsRed. While 2MS2mut-DsRed exhibited the same distribution throughout the cell as BFP, NLS-2MS2mut-DsRed was localized to the nucleus and 2MS2mut-DsRed-NES was localized to the cytoplasm.

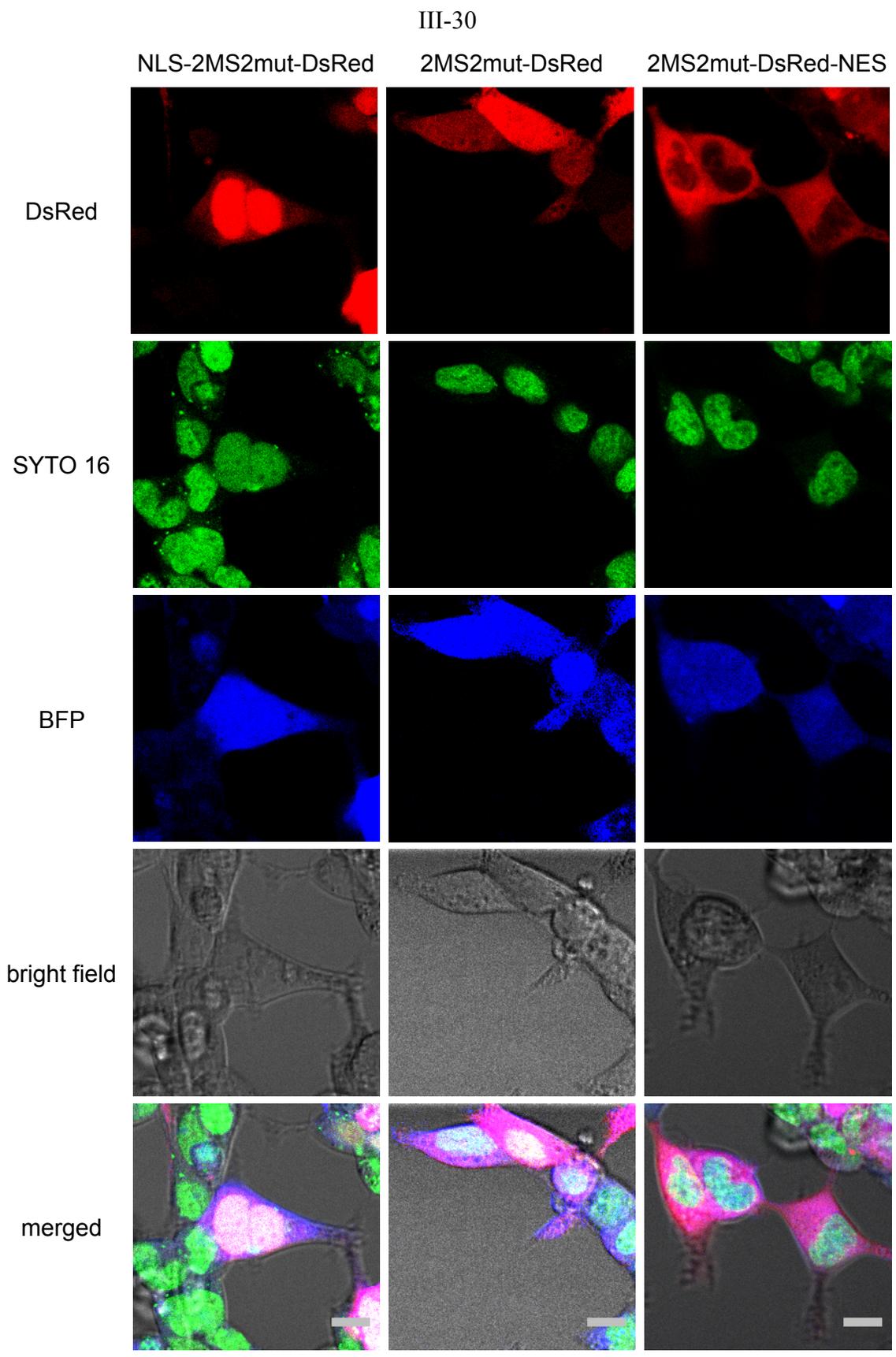


Figure 3.10. Confocal fluorescence microscopy of ligands with and without localization signals: 2MS2mut-DsRed, NLS-2MS2mut-DsRed, 2MS2mut-NES-DsRed. Red, DsRed; blue, BFP; green, SYTO 16 nuclear stain. Scale bars are 10 μm .

We tested the three 2MS2mut-DsRed variants with the best ON (MS2-B11) and OFF (MS2-C1) switch in a stable expression assay. As with 2MS2mut without DsRed, the switches responded to the nuclear-localized ligand to a lesser extent than to either the unlocalized or cytoplasmic-localized ligand, which yielded similar levels of response (MS2-B11: 3.6-fold for NLS, 6.4-fold for unlocalized, 6.5-fold for NES; MS2-C1: 2.8-fold for NLS, 4.6-fold for unlocalized, 4.3-fold for NES) (Figure 3.11). As described above, we assumed that the lower level of switch responsiveness to NLS-2MS2mut-DsRed was likely due to lower steady-state level of protein rather than a specific effect of nuclear localization of ligand.

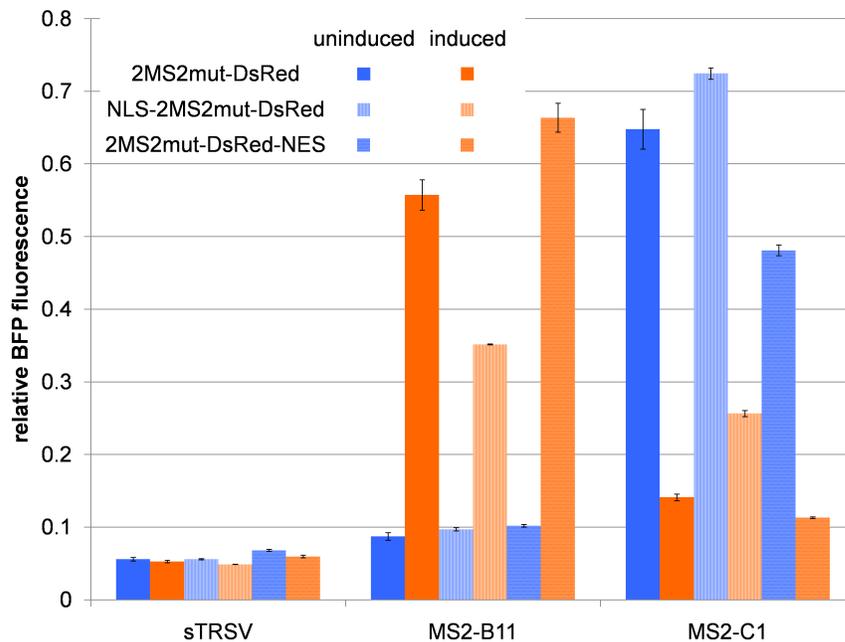


Figure 3.11. Activity of ribozyme switches with improved localized ligands: 2MS2mut-DsRed, NLS-2MS2mut-DsRed, and 2MS2mut-DsRed-NES. Relative BFP fluorescence levels are reported for stably integrated constructs encoding ribozyme switch sequences. Reported values are geometric mean \pm s.d. from biological duplicates and normalized to the non-cleaving sTRSVctrl.

To examine the relationship between ligand expression level and switch response for the three localization variants of 2MS2mut-DsRed, we measured BFP regulation activity over a range of ligand inducer concentrations in a stable expression assay (Figure 3.12A). MS2-B11 and MS2-C1 exhibited a lower response to NLS-2MS2mut-DsRed than the other two ligand variants at all doxycycline concentrations tested. However, comparing BFP regulation activity to DsRed fluorescence (Figure 3.12B) reveals that all three protein variants yield similar switch activity at a given level of DsRed fluorescence. From these results we concluded that switch response is dependent on ligand expression level and not on its localization, and that nuclear and cytoplasmic localization of ligand are each sufficient for switching activity.

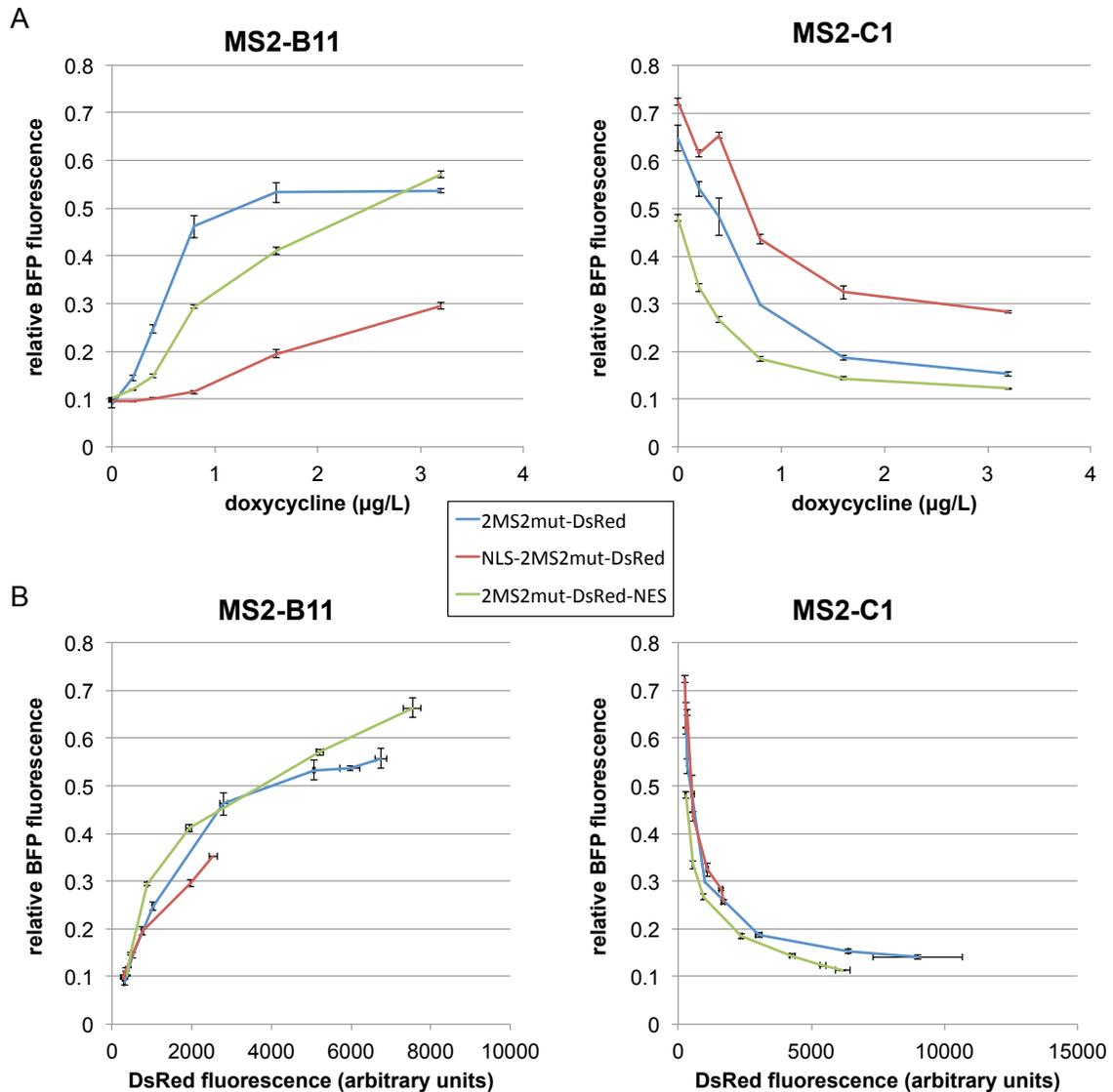


Figure 3.12. Activity of ribozyme switches with improved localized ligands over a range of ligand concentrations. **(A)** Reporter gene expression as a function of doxycycline concentration. Relative BFP fluorescence levels are reported for stably integrated constructs encoding ribozyme switch sequences. Reported values are geometric mean \pm s.d. from biological duplicates and normalized to the non-cleaving sTRSVctrl. **(B)** Reporter gene expression as a function of DsRed fluorescence. The same data sets in B are plotted against DsRed fluorescence levels. Relative BFP and DsRed fluorescence

levels are reported for stably integrated constructs encoding ribozyme switch sequences. Reported values are geometric mean \pm s.d. from biological duplicates and BFP values are normalized to the non-cleaving sTRSVctrl.

Development of ribozyme switch platforms responsive to additional proteins

We attempted to demonstrate the flexibility of our ribozyme switch platform by creating devices responsive to additional proteins. We chose two other bacteriophage proteins with demonstrated sequence-specific RNA binding, *Pseudomonas* phage PP7 coat protein³² and the 1–22 peptide of lambda N protein^{22,33}. We also chose nuclear factor κ B (NF- κ B) and β -catenin, two endogenous proteins involved in transcription and deregulated in many forms of cancer^{7,34–38}. Aptamers to each of these proteins have been selected using SELEX and validated for *in vivo* function^{39–42}.

We designed three ribozyme switches containing the PP7 aptamer (Figure 3.13A). They are based on MS2-C3, which showed a high level of response to MS2, with the PP7 aptamer replacing the MS2 aptamer. The three designs differ from each other in the length of the aptamer stem beyond bulge II. Experiments were performed with transient expression assays using the improved characterization system (Figure 3.2), with the protein ligand expressed from the CMV-TetO₂ promoter. We used a mutant version of PP7 containing amino acid substitutions C68A and C71A to reduce multimerization while maintaining RNA binding affinity⁴³, similar to our strategy with the MS2 ligand. The data demonstrate little gene knockdown and no switching activity for any of the PP7 designs (Figure 3.13B).

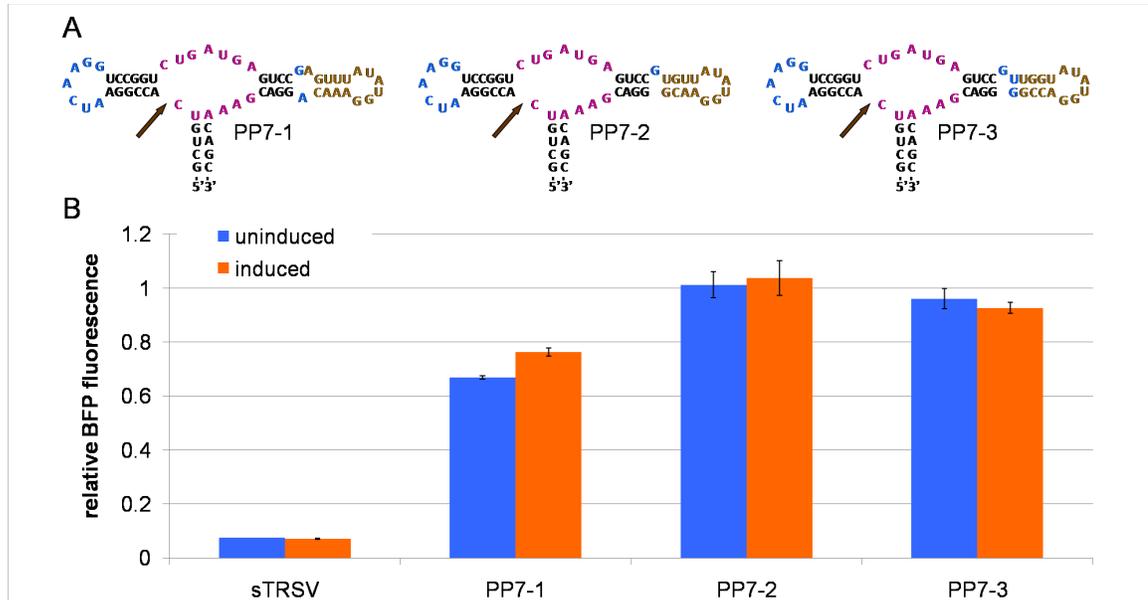


Figure 3.13. Design and testing of PP7-responsive switch designs. (A) Structures of the three designs. Coloring is the same as in Figure 3.1. (B) Activity of PP7-responsive switch designs. Relative BFP fluorescence levels are reported for transiently transfected constructs encoding ribozyme switch sequences cotransfected with a transfection control plasmid. Reported values are geometric mean \pm s.d. from biological duplicates and normalized to the non-cleaving sTRSVctrl.

We designed four ribozyme switches containing the lambda N aptamer (Figure 3.14A). Lambda-1 is based on the previously characterized theophylline-responsive L2b8¹¹ and Lambda-2 is based on MS2-B11, with the lambda N aptamer replacing the theophylline and MS2 aptamers, respectively. The other two designs do not contain a transmitter component, and loop II is replaced by the lambda N aptamer loop in a similar fashion as the MS2-C designs (Figure 3.1). Experiments were performed with transient expression assays using the improved characterization system (Figure 3.2), with the protein ligand expressed from the CMV-TetO₂ promoter. The data demonstrate that the

four ribozyme switch designs exhibit a range of gene knockdown activity; however, none respond to lambda N (Figure 3.14B).

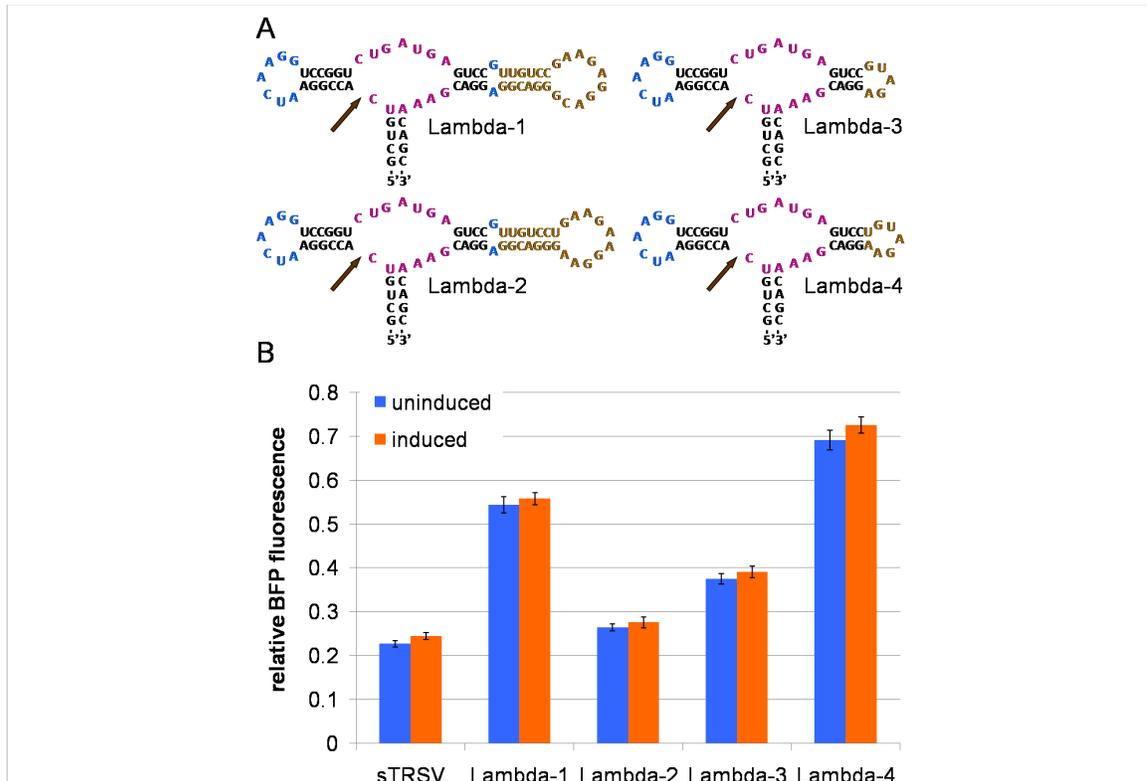


Figure 3.14. Design and testing of lambda-N-responsive switch designs. (A) Structures of the four designs. Coloring is the same as in Figure 3.1. (B) Activity of lambda-N-responsive switch designs. Relative BFP fluorescence levels are reported for transiently transfected constructs encoding ribozyme switch sequences. Reported values are geometric mean \pm s.d. from biological duplicates and normalized to the non-cleaving sTRSVctrl.

We designed thirteen ribozyme switches containing the aptamer for the p50 subunit of NF- κ B and five ribozyme switches containing the aptamer for the p65 subunit

(Figure 3.16A). These designs all contain transmitter components, similar to the MS2-B designs (Figure 3.1), differing from each other in the sequence identity of loop I and the transmitter. For each ligand we designed both ON and OFF switches.

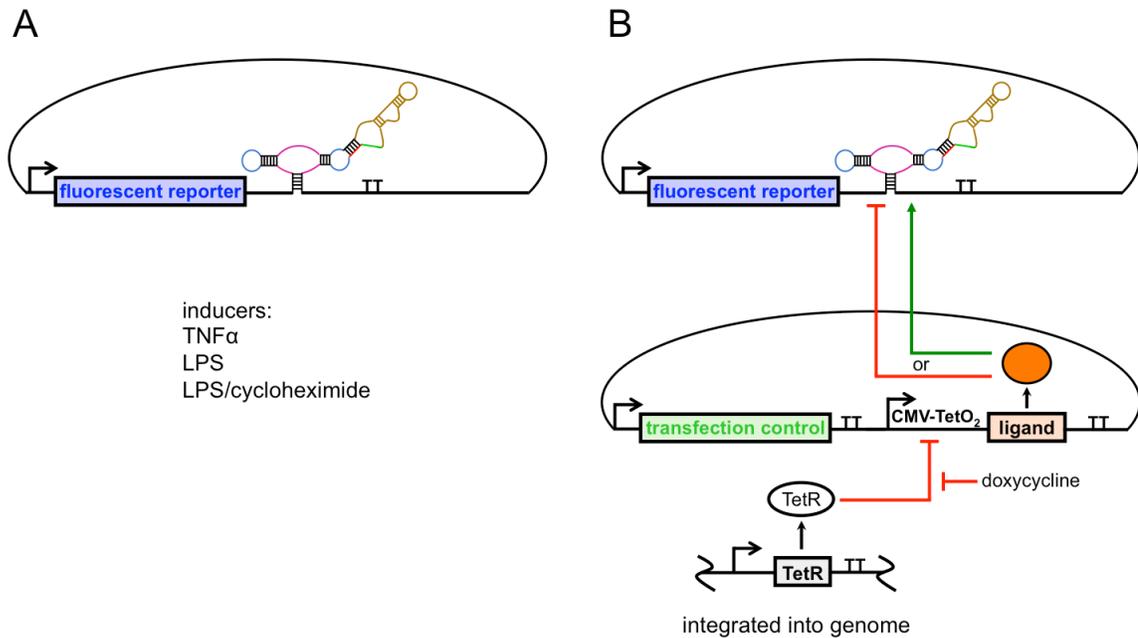


Figure 3.15. NF- κ B-responsive ribozyme switch characterization system. (A) A fluorescent reporter (BFP) with ribozyme switch is encoded on a plasmid, which is transfected into a Flp-In T-REx HEK293 cell line. The NF- κ B signaling pathway is induced with the addition of TNF α , LPS, or LPS in combination with cycloheximide. (B) The switch plasmid is cotransfected with a plasmid encoding the protein ligand (under the control of CMV-TetO₂) and a fluorescent transfection marker. Addition of doxycycline derepresses the CMV-TetO₂ promoter, turning on expression of the ligand, which regulates the activity of ON and OFF switches. Cotransfections were also performed in conjunction with LPS induction.

We employed several different strategies in attempting to elicit switch response to NF- κ B ligand. The NF- κ B p50/p65 heterodimer is normally bound to the inhibitor I κ B, which blocks its NLS, preventing import into the nucleus⁴⁴. Upon activation of the NF- κ B pathway, I κ B is targeted for ubiquitin-dependent degradation, releasing NF- κ B to translocate into the nucleus and activate transcription of its target genes³⁴. It is expected that NF- κ B will not bind to its aptamer when inhibited, as I κ B stabilizes NF- κ B in a conformation with very weak nucleic acid affinity⁴⁵⁻⁴⁷. After transfecting Flp-In T-REx HEK293 cells with plasmids encoding our switches, we activated the NF- κ B pathway with tumor necrosis factor α (TNF α), lipopolysaccharide (LPS), or LPS in combination with cycloheximide (Figure 3.15). We also tried expressing p50 and p65 heterologously, in cotransfections of a plasmid encoding the protein and a plasmid encoding a cognate switch. Finally, we tried heterologous protein expression in combination with LPS induction. Although the designs exhibited a range of gene knockdown activity, none of them displayed switching activity under any of the conditions tested (Figure 3.16B).

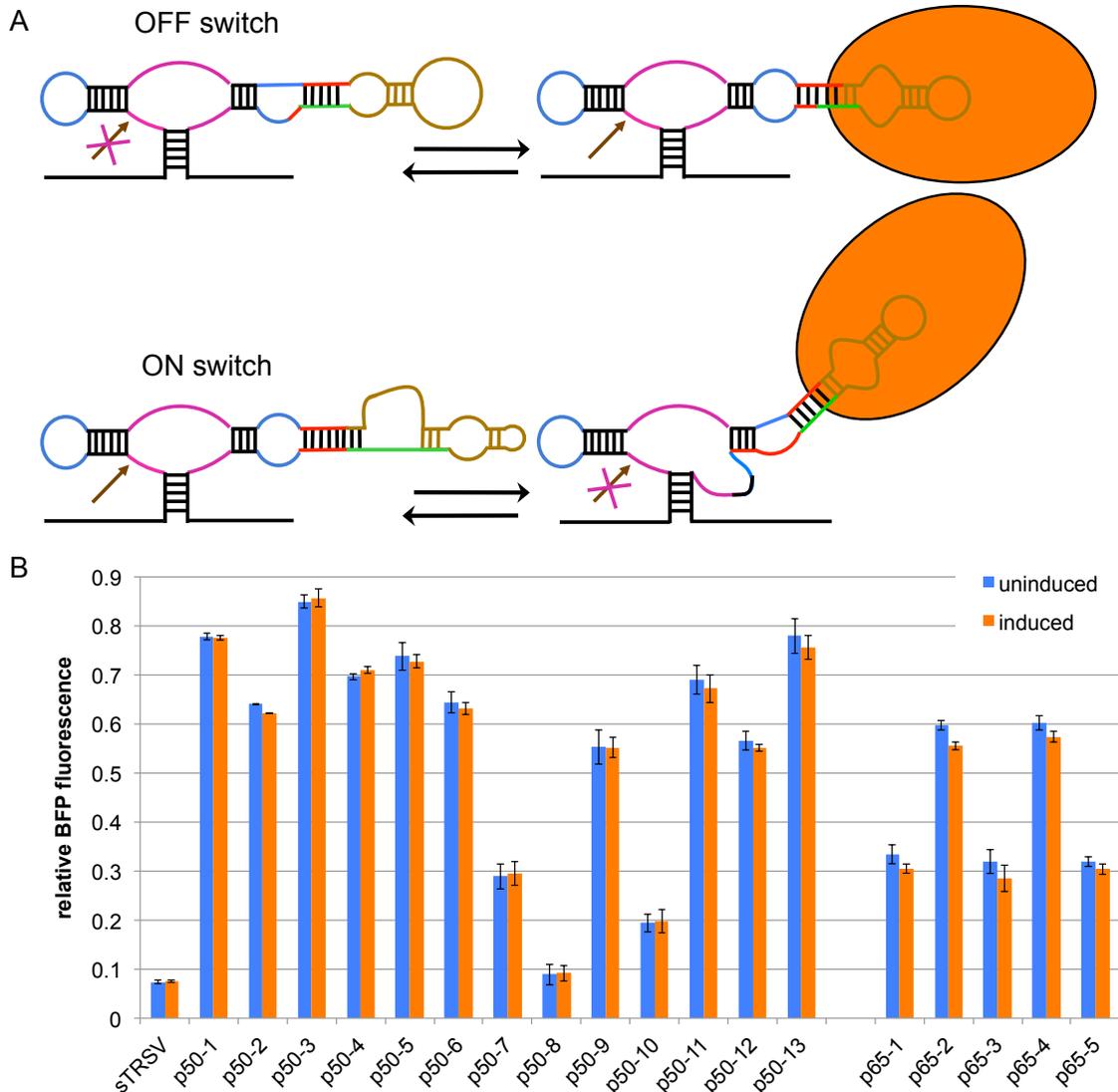
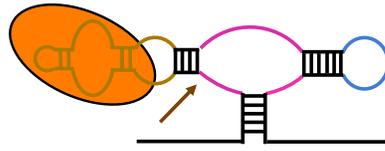


Figure 3.16. Design and testing of NF- κ B-responsive switch designs. **(A)** ON and OFF switches were designed with aptamers for p50 and p65 subunits of NF- κ B. Coloring is the same as in Figure 3.1. See Supplementary Table 3.1 for sequences. **(B)** Activity of NF- κ B-responsive switch designs. Relative BFP fluorescence levels are reported for transiently transfected constructs encoding ribozyme switch sequences cotransfected with a plasmid encoding p50 or p65 and a fluorescent transfection control. Reported values

are geometric mean \pm s.d. from biological duplicates and normalized to the non-cleaving sTRSVctrl.

We designed two sets of ribozyme switches containing the β -catenin aptamer (Figure 3.17A). In one set of designs (Bcat-A) the β -catenin aptamer replaced loop I of four different hammerhead ribozymes similar in sequence and structure to sTRSV (sLTSV⁻, sLTSV⁺, CChMVd⁻, and SCMoV⁺), following a similar strategy as the MS2-C designs. The other set of designs (Bcat-B) was similar to the MS2-responsive sequential aptamer and ribozyme designs (MS2-D), with the β -catenin aptamer just upstream of the ribozyme. These six designs varied in the length and sequence identity of the transmitter component.

A
 III-41
 Bcat-A



Bcat-B

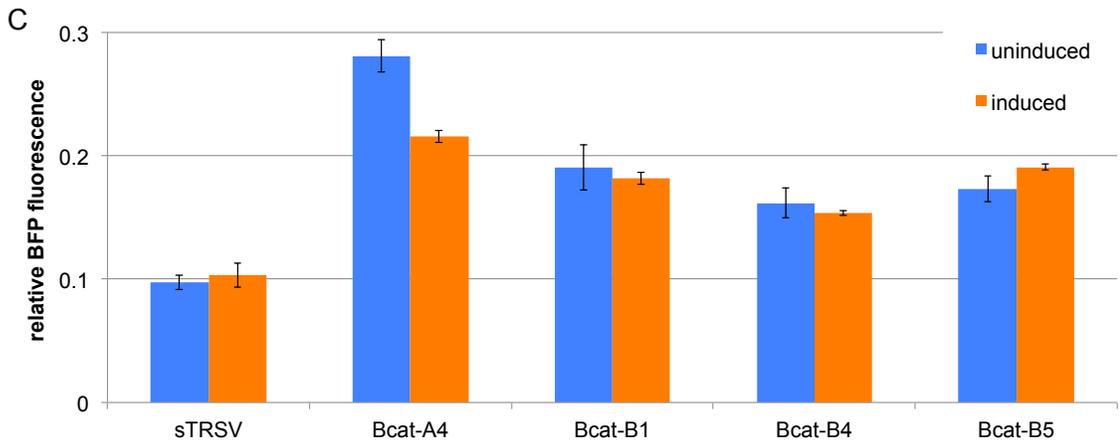
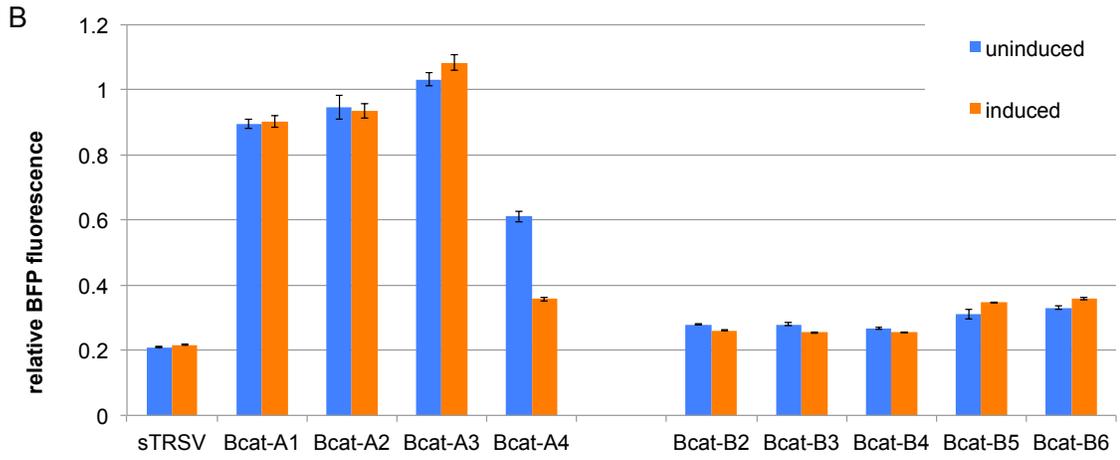
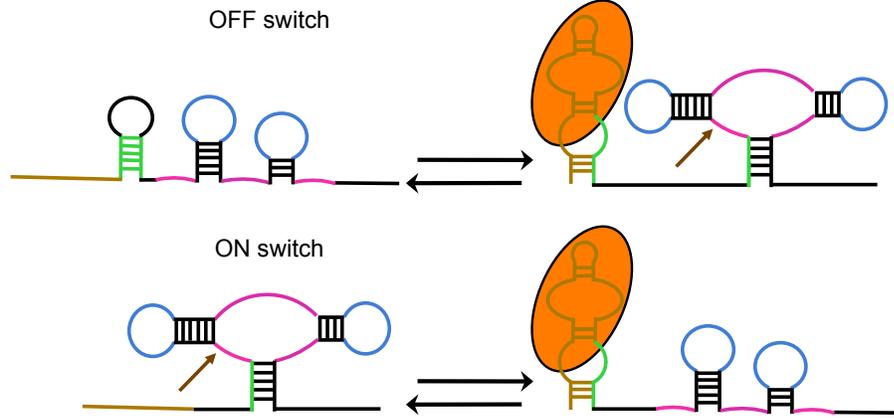


Figure 3.17. Design and testing of β -catenin-responsive switch designs. **(A)** β -catenin-responsive loop I replacement (Bcat-A) and sequential aptamer and ribozyme (Bcat-B) switch designs. Coloring is the same as in Figure 3.1. See Supplementary Table 3.1 for sequences. **(B)** Activity of β -catenin-responsive switch designs in a transient expression assay. Relative BFP fluorescence levels are reported for transiently transfected constructs encoding ribozyme switch sequences. Reported values are geometric mean \pm s.d. from biological duplicates and normalized to the non-cleaving sTRSVctrl. **(C)** Activity of β -catenin-responsive switch designs in a transient expression assay with transfection control. Relative BFP fluorescence levels are reported for transiently transfected constructs encoding ribozyme switch sequences cotransfected with a transfection control plasmid. Reported values are geometric mean \pm s.d. from biological duplicates and normalized to the non-cleaving sTRSVctrl.

β -catenin translocates into the nucleus upon activation of the Wnt signaling pathway, and in the absence of signaling is phosphorylated and degraded by the proteasome³⁸. After transfecting Flp-In T-REx HEK293 cells with plasmids encoding our switches, we activated the pathway with the addition of the cytokine Wnt3A to the cell culture media but observed no switch response. We also tried heterologously expressing the Arm 1–12 domain of β -catenin⁴⁸, which has been shown to have higher affinity for the aptamer than the full-length protein⁴⁹, from the CMV-TetO₂ promoter in a transient transfection assay, as we did with the MS2 proteins (Figure 3.2). Most of the loop I replacement designs (Bcat-A) exhibited little gene knockdown activity, while the sequential designs (Bcat-B) exhibited high levels of gene knockdown activity, similar to

the MS2-D designs (Figure 3.17B). Only one design, Bcat-A4, exhibited responsiveness to heterologous β -catenin in this experimental system. Induction with Wnt3A in addition to the heterologous expression of β -catenin had no effect on switch activity. We reexamined four switch designs in an assay with a transfection control plasmid and Bcat-A4 exhibited 1.3-fold OFF-switching (Figure 3.17C). Stable integration of the Bcat-A4 and heterologous β -catenin construct into Flp-In T-REx HEK293 cells resulted in no switch response, likely due to the ligand being expressed at a lower level than in transient transfections, as described above (Figure 3.5).

Discussion

We developed protein-responsive gene-regulatory devices based on the ribozyme switch platform previously developed in the Smolke laboratory¹¹. We first screened a large number of designs with different architectures, and found that directed conformational changes of ribozyme structure are important for ligand responsiveness. Specifically, coupling the MS2 aptamer directly to a ribozyme loop, without a transmitter component (MS2-A), resulted in devices that did not respond to MS2 ligand. Designs containing a transmitter that altered the structure of the stem and catalytic core (MS2-B), analogous to previously developed ribozyme switches¹¹, were among the most highly functional switch designs tested. The designs containing a transmitter that altered the structure of the stem and loop (MS2-C) also led to functional switches. Three transmitter designs, MS2-B1, MS2-B3, and MS2-B7, were altered by replacing loop I with an alternative sequence, generating MS2-B2, MS2-B10, and MS2-B11. In all three cases

the altered loop resulted in improved gene knockdown and switching activity (Figure 3.7B). The sequential designs, with the aptamer immediately upstream of the ribozyme in the transcript (MS2-D), were mostly insensitive to ligand but one (MS2-D5) exhibited a moderate amount of switching activity (up to 2.4-fold) and the lowest basal expression level of all designs tested, with gene knockdown activity comparable to wild-type ribozyme (sTRSV).

We further improved our characterization system by optimizing the MS2 ligand. Responsiveness to wild-type MS2 in our initial switch characterization study was at most 1.8-fold (Figure 3.4, MS2-B2). Replacing wild-type MS2 with a mutant containing two amino acid substitutions that prevent multimerization²⁷ improved sensitivity for most of the switches tested (Figure 3.7A). The fused dimer of this mutant MS2 elicited an even greater switch response, including from MS2-B1, which was unresponsive to the wild-type and mutant monomer (Figure 3.7A). We concluded that the fused dimer of mutant MS2 was the best ligand for characterizing our switch designs.

Using our optimized ligand we were able to demonstrate high levels of ON-switching with MS2-B11 (up to 4.1-fold) and OFF-switching with MS2-C1 (up to 3.8-fold). In contrast, previously described small-molecule-responsive ribozyme switches from the Smolke laboratory^{15,17} have exhibited at most 2.1-fold ON switching (L2b18tc, tetracycline-responsive) and 1.7-fold OFF-switching (Lb2OFF, theophylline-responsive) in human cell lines. Switching activity was improved by incorporating multiple copies of the ribozyme switch into the 3' UTR of the target gene, resulting in up to 3.5-fold ON-switching (L2b8, 2 copies and L2b9, 3 copies)¹⁷. However, OFF switch activity has not been improved by this method.

We demonstrated the localization requirements of our ribozyme switches by assaying for sensitivity to our optimized ligand with fused localization signals. The initial implementation of these localization signals did not fully direct the protein to the desired subcellular compartment (Figure 3.8). Although the NES-tagged protein was present in the nucleus at levels no higher than a cytoplasmic control protein, the NLS-tagged protein was present in both the nucleus and the cytoplasm. We hypothesized that despite the fusion with localization signals the protein was passively diffusing through the nuclear pore due to its small size. We therefore created fusion proteins with DsRed, whose increased size was predicted to prevent passive diffusion through the nuclear pore and result in full localization, which was verified with confocal fluorescence microscopy (Figure 3.10).

Testing of localized variants of 2MS2mut-DsRed with MS2-B11 and MS2-C1 revealed little dependence of switch activity on ligand localization. Unlocalized and cytoplasmic-localized ligands elicited similar levels of response for each switch, while the nuclear-localized ligand elicited much lower levels of response (Figure 3.11). We have observed by flow cytometry that NLS-tagged fluorescent proteins generate fluorescence levels less than half those of analogous untagged proteins (Figure 3.12B and Ryan Bloom, unpublished results). It is possible that the addition of the seven-amino-acid NLS to the N-terminus of a fluorescent protein affects its three-dimensional structure in a way that negatively impacts its fluorescent output; however, we believe this is unlikely. Immunoblotting indicates that the total amount of MS2 protein per cell is lower with the NLS tag than without (Figure 3.8A), suggesting that the lower level of fluorescence observed from NLS-tagged proteins is due to lower steady-state protein

levels and not to a decrease in fluorescence output per molecule. This conclusion is further supported by our finding that switch response is dependent on ligand expression level as measured by fluorescence (Figure 3.12B). Assuming that each of the three localization variants of 2MS2mut-DsRed exhibits roughly the same fluorescence output per molecule, the data show that switch response is correlated with ligand expression level regardless of localization.

Nuclear and cytoplasmic localization of ligand each appear to be sufficient for switching activity in human cells. This is somewhat surprising for an ON switch such as MS2-B11, which cleaves in the absence of ligand. One might expect this switch to cleave during or immediately after transcription if ligand is absent from the nucleus, or to cleave after nuclear export if ligand is absent from the cytoplasm. However, our data did not support these initial expectations.

Although our confocal fluorescence microscopy measurements show clear localization, we cannot rule out the possibility that a small amount of the NLS-tagged ligand is found in the cytoplasm and a small amount of the NES-tagged ligand is found in the nucleus. The presence of such mislocalized protein could prevent precise examination of the effect of ligand localization on ribozyme switch activity if that small amount were sufficient to prevent ribozyme ON switch cleavage. However, the correlation between switch response and DsRed fluorescence (Figure 3.12B) suggests that this is unlikely. If, for example, nuclear localization were necessary for switch activity, and switches appeared to respond to NES-tagged ligand, we would expect the required amount of NES-tagged ligand expression to be much higher than NLS-tagged ligand to yield the same level of switch activity, as most of the NES-tagged protein would

be in the cytoplasm and unable to affect the switch. We observed that similar ligand expression levels, as measured by fluorescence, resulted in similar levels of switching regardless of localization signal.

Our data suggest that the ribozyme switches do not cleave before nuclear export to a significant degree, nor do they cleave in the cytoplasm to a significant degree when the ligand is present in the nucleus. We speculate that ribozyme cleavage in the nucleus is low, possibly due to prevention of proper folding by binding of proteins that form the messenger ribonucleoprotein (mRNP). This would minimize ribozyme cleavage before the switch is exposed to cytoplasmic-localized ligand. When ligand is localized to the nucleus, we speculate that it is able to bind to the aptamer and is carried out of the nucleus by the mRNA during nuclear export. The ability of mRNA containing an aptamer to carry MS2 out of the nucleus has been previously demonstrated²⁰, but this behavior may depend on the small size of the MS2 protein and not be generalizable to larger proteins. After export to the cytoplasm, dissociation of the ligand from the ribozyme switch would be favored due to the extremely low local concentration of free ligand. A low ligand off-rate and a low ribozyme cleavage rate would both contribute to allowing the mRNA to be translated before cleavage. Functional MS2-responsive ribozyme switches presented here exhibit off rate constants in the range of 0.001–0.05 s⁻¹ (Andrew Kennedy, unpublished results) and cleavage rate constants in the range of 0.05–1 min⁻¹ (Andrew Kennedy, unpublished results). ON switches with higher cleavage rates would be expected to be less responsive to localized ligand (nuclear or cytoplasmic), and ON switches with higher ligand off-rates would be expected to be less responsive to nuclear-localized ligand.

In the future the ribozyme switch mechanism of action could be more thoroughly investigated. Studies using switches with a range of cleavage-rate constants and ligand off-rate constants would establish the relationship between these parameters and *in vivo* response to localized ligands. This in turn would confirm or disprove our supposition that nuclear and cytoplasmic localization are each sufficient for switching activity. If it is indeed true that ligand localization does not negatively impact switching activity, then our ribozyme switch platform would be capable of sensing proteins in the nucleus or the cytoplasm. In contrast, switches based on regulation of shRNA processing⁹ or splicing¹⁰ are limited to sensing cytoplasmic or nuclear proteins, respectively.

We attempted to develop ribozyme switches responsive to additional proteins. We suspected that generating switches with aptamers for bacteriophage proteins (PP7 and lambda N) would be straightforward, basing our designs on functional MS2-responsive switches. However, none of these designs responded to ligand. We next attempted to generate ribozyme switches responsive to NF- κ B and β -catenin, two proteins involved in transcription and deregulated in many forms of cancer^{7,34-37}, for which there exist *in vitro* selected aptamers with validated *in vivo* function³⁹⁻⁴². Although NF- κ B and β -catenin are endogenously expressed in our human cell line, we did not expect them to be available for binding to ribozyme switches without activation of their signaling pathways or heterologous overexpression. We were unable to demonstrate NF- κ B-responsiveness, but one β -catenin switch, Bcat-A4, exhibited 1.3-fold OFF-switching in response to transient heterologous expression of β -catenin. We suspect that Bcat-A4 did not respond to endogenous or stably expressed β -catenin because the steady-state levels of protein were too low. In the future high-throughput *in vivo* screening methods⁵⁰⁻⁵² could be used

to assay large libraries of devices to explore a wider design space and achieve greater success generating functional switches responsive to new protein ligands.

The protein-responsive ribozyme switch platform we have developed is unique in its ability to respond to ligands in both the nucleus and the cytoplasm, while previously reported protein-responsive switches can function in only one compartment^{9,10,53,54}. One potential limitation of our platform is that it is not able to detect changes in protein localization. However, our platform's capability of sensing both nuclear and cytoplasmic proteins may be an important advantage for its use as a noninvasive reporter or phenotypic controller in future applications.

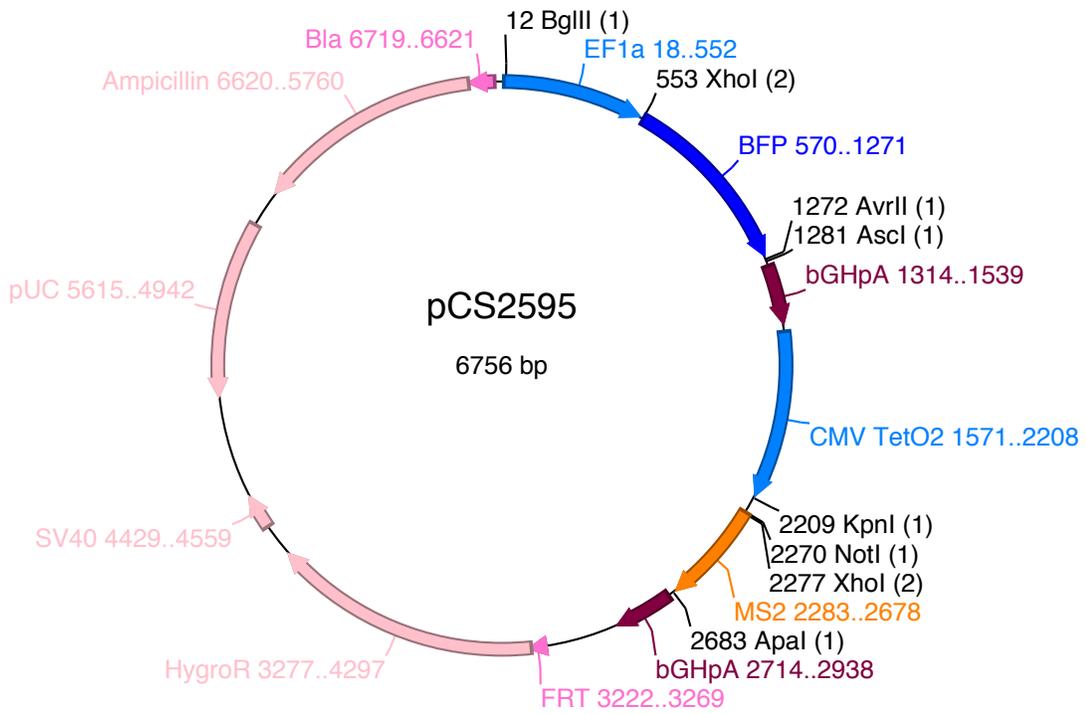
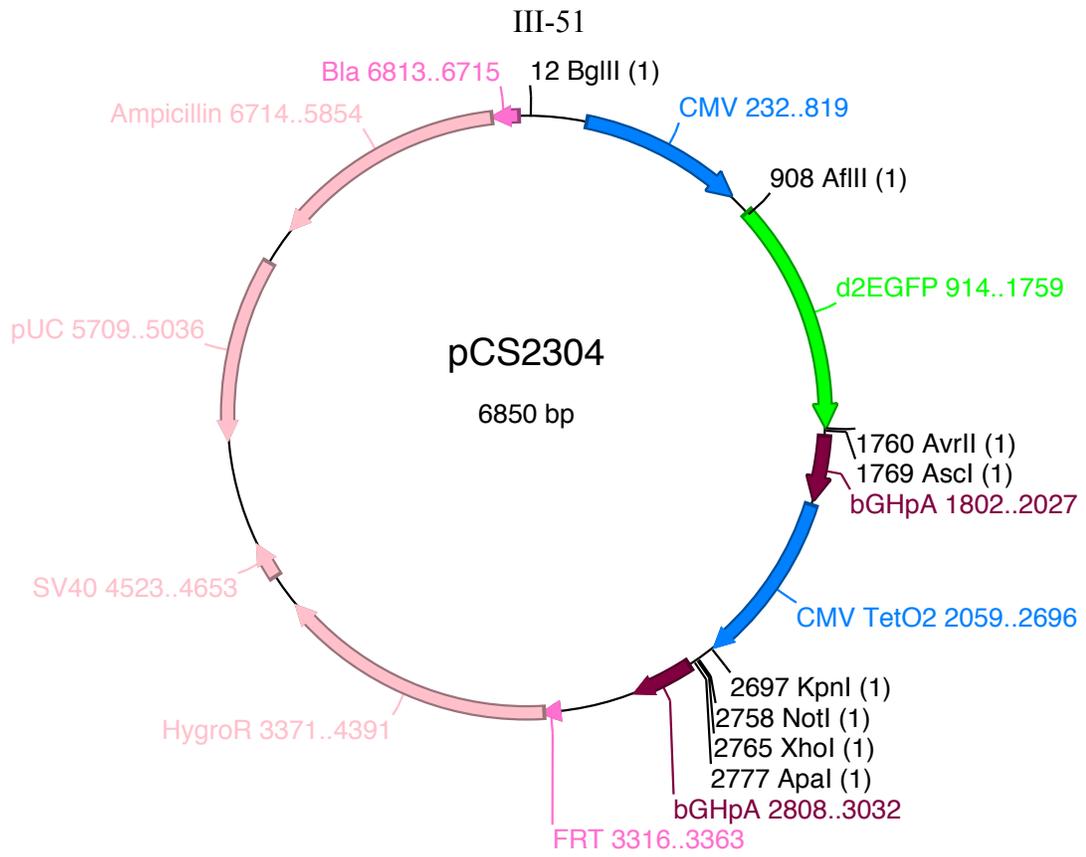
Methods

Plasmid construction

All plasmids were constructed using standard molecular biology techniques. Oligonucleotides were synthesized by Integrated DNA Technologies and the Stanford Protein and Nucleic Acid Facility. Cloning enzymes, including restriction enzymes and T4 DNA ligase, were obtained from New England Biolabs. Ligation products were electroporated into *Escherichia coli* DH10B (Life Technologies) using a GenePulser XP (Bio-Rad Laboratories) system or transformed into *E. coli* One Shot Top 10 (Life Technologies) using standard methods. Clones were screened using colony polymerase chain reaction (PCR) and verified by sequencing (Elim Biopharmaceuticals). 15% glycerol stocks were made from *E. coli* in logarithmic growth phase and stored at -80°C .

III-50

A standardized cloning method was developed to facilitate insertion of ligand-responsive devices and ligand coding regions into a single plasmid backbone. A DNA fragment encoding d2EGFP with a bGHpA signal and the CMV-TetO2 promoter was synthesized by GeneArt (Life Technologies) and inserted into pcDNA5/FRT (Life Technologies) between the restriction sites AflII/KpnI to form pCS2304 (Figure 3.18), which contained a CMV promoter expressing d2EGFP and FRT recombinase sites compatible with stable integration into the genome of Flp-In T-REx HEK293 cells (Life Technologies) to create isogenic stable cell lines. The coding region of the fusion protein MS2-DsRedMonomer was PCR amplified from pCS1392 (courtesy Stephanie Culler) using the primers No NLS A/X Fwd and DsRed A/X Rev and inserted into pCS2304 between XhoI/ApaI to form pCS2359. MS2-responsive ribozyme switch designs were inserted into pCS2359 between AvrII/AscI.



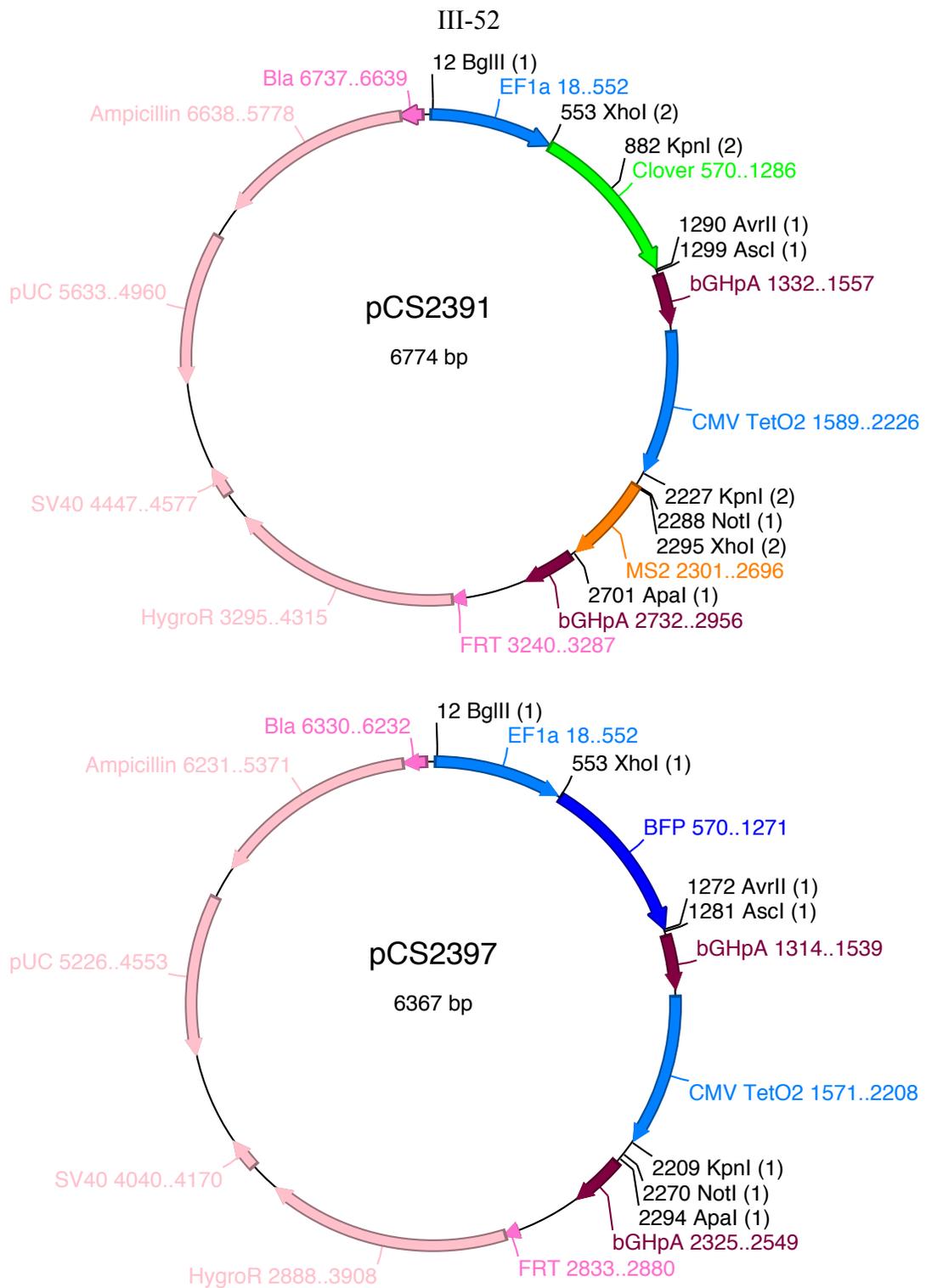


Figure 3.18. Plasmid maps.

III-53

An Efl α promoter with the coding region of BFP was PCR amplified from pCS2585 (courtesy Melina Mathur) using the primers EF1BFP Fwd and EF1BFP Rev and inserted between BglII/AvrII, and the coding region of MS2 was PCR amplified from pCS1392 using No NLS A/X Fwd and MS2 A/X Rev and inserted between XhoI/ApaI into pCS2304 to form pCS2595 (Figure 3.18). MS2-responsive ribozyme switch designs were inserted into pCS2595 between AvrII/AscI.

The plasmid pCS2359 was digested with NheI/AvrII to remove d2EGFP and ligated to form pCS2406. The coding region of MS2 was PCR amplified from pCS1392 using the primers No NLS A/X Fwd and MS2 A/X Rev and inserted into pCS2406 between XhoI/ApaI to form pCS2409.

The coding region of MS2mut (V75E/A81G) was PCR amplified from the MBP-MS2-His plasmid (courtesy Rachel Green, Department of Molecular Biology and Genetics, Johns Hopkins University) using the primers MS2 NotI F and MS2 ApaI R and inserted into pCS2595 between NotI/ApaI to form pCS2631. MS2-responsive ribozyme switch designs were inserted into pCS2631 between AvrII/AscI.

A DNA fragment encoding 2MS2mut (MS2 V75E/A81G head-to-tail fused dimer) was synthesized by GeneArt (Life Technologies) and inserted into pCS2595 between NotI/ApaI to form pCS2686. The coding region of 2MS2mut was PCR amplified from pCS2686 using the primers NLS MS2 F and 2MS2mut R to add an N-terminal NLS, and using the primers 2MS2mut F and NES MS2 R to add a C-terminal NES, and inserted into pCS2595 between NotI/ApaI to form pCS2747 and pCS2787, respectively. MS2-responsive ribozyme switch designs were inserted into pCS2686, pCS2747, and pCS2787 between AvrII/AscI.

The coding region of DsRedMonomer was PCR amplified from pCS2359 using the primers DsRed GF and DsRed GR, and using the primers DsRed GF and DsRed NES R to add a C-terminal NES. The resulting DNA fragments were inserted into plasmids digested with *ApaI* using Gibson assembly⁵⁵ as follows: DsRed into pCS2686 to form pCS2897, DsRed into pCS2747 to form pCS2902, and DsRed-NES into pCS2686 to form pCS2907. MS2-responsive ribozyme switch designs were inserted into pCS2897, pCS2902, and pCS2907 between *AvrII/AscI*.

The coding region of Clover GFP was PCR amplified from pCS2586 (courtesy Melina Mathur) using the primers GF Clover and GR Clover, and pCS2595 was PCR amplified using the primers GF EF1 and GR EF1. The resulting DNA fragments were assembled using Gibson assembly⁵⁵ to form pCS2391 (Figure 3.18), which contained Clover GFP in place of BFP.

The coding region of the PP7 coat protein was PCR amplified from Addgene plasmid 28174 (Kathleen Collins) using the primers PP7 *NotI* F and PP7 *ApaI* R and inserted into pCS2595 between *NotI/ApaI* to form pCS2847. PP7-responsive ribozyme switch designs were inserted into pCS2847 between *AvrII/AscI*.

The DNA fragment lambda N was inserted into pCS2595 between *NotI/ApaI* to form pCS2816. Lambda-N-responsive ribozyme switch designs were inserted into pCS2816 between *AvrII/AscI*.

The DNA fragment insertNA was inserted into pCS2595 between *NotI/ApaI* to form pCS2397 (Figure 3.18). NF- κ B-responsive ribozyme switch designs were inserted into pCS2397 between *AvrII/AscI*. The coding region of NF- κ B p50 was PCR amplified from pCS1806 (courtesy Stephanie Culler) using the primers p50 *NotI* F and p50 *ApaI* R

and inserted into pCS2391 between NotI/ApaI to form pCS2604. The coding region of NF- κ B p65 was PCR amplified from pJ1448 (courtesy Louis Maher III, Department of Biochemistry and Molecular Biology, Mayo Clinic College of Medicine) using the primers p65 NotI F and p65 ApaI R and inserted into pCS2391 between NotI/ApaI to form pCS2605.

β -catenin-responsive ribozyme switch designs were inserted into pCS2397 between AvrII/AscI. The coding region of β -catenin Arm 1–12 was PCR amplified from Addgene plasmid 17198 (Randall Moon) using the primers b-cat NotI F and b-cat ApaI R and inserted into pCS2595 between NotI/ApaI to form pCS2824. β -catenin-responsive ribozyme switch designs were inserted into pCS2824 between AvrII/AscI.

Human cell culture

Flp-In T-REx HEK293 cells (Life Technologies) were cultured in 10 mL (10 cm dish) or 3 mL (6 cm dish) Dulbecco's Modified Eagle's medium (DMEM) (Life Technologies) supplemented with 10% fetal bovine serum (FBS) (Life Technologies), 100 mg/L zeocin (Life Technologies), and 5 mg/L blasticidin (Life Technologies) in a humidified incubator at 37°C and 5% CO₂. Cells were seeded at 2x10⁴ cells/mL and passaged regularly using 0.25% trypsin-EDTA (Life Technologies), with media replaced every 48–72 hours. Cells stably integrated with Flp-In constructs were cultured similarly, except the cell culture media were supplemented with 100 mg/L hygromycin B (Life Technologies) and no zeocin.

Stable cell line generation

Flp-In T-REx HEK293 cells were seeded at 1×10^5 cells/mL in 2 mL (6-well plate) DMEM with 10% FBS. 24 hours later the cells were cotransfected with a pcDNA5/FRT-derived plasmid and pOG44 (Life Technologies) in a 1:9 ratio using FuGENE HD (Promega) according to the manufacturer's instructions. Typically DNA and FuGENE were incubated together in Opti-MEM in a 1:3:50 (g:L:L) ratio for approximately 1 hour, with 2 mL samples receiving 2 μ g of DNA. 24 hours after transfection the cells were resuspended using 0.25% trypsin-EDTA and DMEM with 10% FBS, and $\frac{1}{4}$ of the cells were used to seed 2 mL (6-well plate) DMEM with 10% FBS. 24 hours later the media were replaced with DMEM with 10% FBS, 200 mg/L hygromycin B, and 5 mg/L blasticidin. The media were replaced every 72–96 hours until macroscopic colonies were visible, usually after 10–14 days. Colonies were pooled together with 0.25% trypsin-EDTA and passaged into DMEM with 10% FBS, 100 mg/L hygromycin B, and 5 mg/L blasticidin. 10% dimethyl sulfoxide (DMSO) stocks were made from resuspended cells, cooled by 1 degree/minute to -80°C , then stored at -320°C .

Transient transfection

Flp-In T-REx HEK293 cells were seeded at 1×10^5 cells/mL in 500 μ L (24-well plate), 10 mL (10 cm dish), or 400 μ L (8-chambered coverglass) DMEM with 10% FBS. 21–27 hours (flow cytometry assay) or 48 hours (cellular fractionation and extraction) or 24 hours (confocal microscopy) after seeding the cells were transfected with one or two plasmids using FuGENE HD (Promega) or Lipofectamine 2000 (Life Technologies) according to the manufacturer's instructions. Typically DNA and FuGENE HD were

incubated together in Opti-MEM in a 1:3:50 (g:L:L) ratio for 1 approximately hour, while DNA and Lipofectamine were incubated together in Opti-MEM in a 1:2:100 (g:L:L) ratio for 5 minutes. With either transfection reagent, 500 μ L samples received 500 ng of DNA, 10 mL samples received 10 μ g of DNA, and 400 μ L samples received 400 ng of DNA.

Flow cytometry

18–28 hours after seeding (30–75 minutes after transfection if applicable) doxycycline was added to derepress the CMV-TetO₂ promoter. 24–26 hours after transfection with plasmids encoding NF- κ B- or β -catenin-responsive ribozyme switch designs inducer molecules were added. 10 μ g/L TNF α (Sigma-Aldrich), 500 μ g/L LPS (Sigma-Aldrich), or 500 μ g/L LPS in combination with 10 mg/L cycloheximide (Sigma-Aldrich) were used to induce the NF- κ B pathway, and 200 μ g/L Wnt3A (R&D Systems) was used to induce the Wnt pathway. 42-52 hours after transfection fluorescence data were obtained by flow cytometry using the MACSQuant VYB equipped with 405 nm, 488 nm, and 561 nm lasers (Miltenyi Biotec). Viability was gated by side scatter and electronic volume, and viable cells were further gated for either DsRed, GFP, or BFP expression, which served as transfection controls. DsRed, GFP, and BFP fluorescence was measured through 615/20 nm, 525/50 nm, and 450/50 nm band-pass filters, respectively. Data were analyzed using FlowJo (Tree Star Inc.). Geometric mean values from biological replicates were reported with an error range of ± 1 standard deviation. Geometric mean fluorescence values were normalized to those of a control with no ribozyme or the inactive ribozyme sTRSVctrl.

Cellular fractionation and extraction

1–1.5 hours before transfection cell culture media were replaced with media containing 1 mg/L doxycycline to derepress the CMV-TetO₂ promoter. 50 hours later cytoplasmic and nuclear extracts were prepared using the CelLytic NuCLEAR Extraction Kit (Sigma) with isotonic lysis buffer and IGEPAL CA-630 detergent according to the manufacturer's instructions, although modifications were made to the protocol to minimize cross-contamination between the nuclear and cytoplasmic fractions. Briefly, cells were washed with PBS, scraped off of the culture dishes, and centrifuged at 500 x g for 5 minutes. Packed cells were resuspended in isotonic lysis buffer and incubated for 1 minute on ice, then incubated with IGEPAL CA-630 at a final concentration of 0.04% on ice for 3 minutes. Lysed cells were centrifuged at 5000 x g for 30 seconds and the cytoplasmic fraction (supernatant) was collected. The pelleted nuclei were washed with isotonic lysis buffer and 0.04% IGEPAL CA-630, then centrifuged at 5000 x g for 30 seconds. The pelleted nuclei were resuspended in extraction buffer and agitated at 4°C for 30 minutes. Lysed nuclei were centrifuged at 18000 x g for 10 minutes and the nuclear fraction (supernatant) was collected.

Immunoblotting

A standard Bradford assay using Protein Assay reagent (Bio-Rad Laboratories) was performed with a BSA standard to determine protein concentrations. Samples were run on NuPAGE 4–12% Bis-Tris Gels (Life Technologies) in NuPAGE MOPS buffer (Life Technologies) at 150 V for 1 hour. Transfer was performed with extra thick blot paper (Bio-Rad Laboratories) and 0.45 µm Protran nitrocellulose transfer membrane

(Whatman) in 2x NuPAGE transfer buffer (Life Technologies) and 20% methanol using a Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad Laboratories) at 15 V for 15 minutes. Membranes were blocked with TBST (20 mM Tris, 137 mM sodium chloride, 0.1% Tween-20) and 5% bovine serum albumin (BSA) (Calbiochem) at room temperature for 1 hour, then rinsed with TBST twice for 5 minutes each. Membranes were probed with rabbit anti-enterobacteriophage MS2 coat protein, anti-GAPDH, and anti-HDAC1 polyclonal antibodies (Millipore) in TBST and 1% BSA at 4°C for 16 hours, then rinsed with TBST twice for 5 minutes each. Membranes were probed with sheep anti-rabbit IgG, horseradish peroxidase (HRP) conjugate polyclonal antibody (Millipore) at room temperature for 1 hour, then rinsed with TBST twice for 5 minutes each. HRP signal was detected using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) according to the manufacturer's instructions and a G:Box Chemi XT4 imaging system (Syngene). Band intensity was calculated with GeneTools software (Syngene).

Immunofluorescence microscopy

Flp-In T-REx HEK293 cells with stably integrated 2MS2mut constructs were seeded at 1×10^5 cells/mL in 1 mL (4-chambered coverglass) DMEM with 10% FBS, 100 mg/L hygromycin B, and 1 mg/L doxycycline to derepress the CMV-TetO₂ promoter. Approximately 43 hours after seeding cells were washed with phosphate buffered saline (PBS) (Life Technologies) and fixed for 15–20 minutes using HistoChoice MB tissue fixative (AMRESCO). Cells were washed twice with PBS, blocked with PBS and 1.5% BSA for 1 hour, and washed with PBS for 5 minutes. Cells were probed with rabbit anti-enterobacteriophage MS2 coat protein at 4°C for approximately 17 hours, then washed

with PBS for 5 minutes. Cells were probed with sheep anti-rabbit fluorescein conjugated [F(ab')₂ fragments] polyclonal antibody (Chemicon) for 30 minutes, then washed thrice with PBS for 5 minutes each. Cell nuclei were counterstained using 250 µg/L 7-AAD (Life Technologies) in PBS for 5 minutes, then washed with PBS. Cells were imaged on a Zeiss Axiovert 200M fluorescence microscope (Zeiss) with a 20x objective using the AxioVision software (Zeiss). Images were exported and brightness and contrast were adjusted using FIJI.

Confocal fluorescence microscopy

Approximately 30 minutes after transfection 1 mg/L doxycycline was added to derepress the CMV-TetO₂ promoter, and 23 hours later media were replaced with media containing 1 mg/L doxycycline. 24 hours later cell nuclei were counterstained using 250 nM SYTO 16 (Life Technologies). 24 hours after counterstaining cells were imaged on a Leica TCS SP8 confocal microscope (Leica Microsystems) with a 20x objective using the Leica Application Suite Advanced Fluorescence software (Leica Microsystems). Images were exported and brightness and contrast were adjusted using FIJI.

Acknowledgements

We thank Andrew Kennedy and Leopold d'Espaux for designing ribozyme switches, designing experiments, and evaluating results; Marty Bigos for help evaluating flow cytometry results; Kitty Lee and the Stanford Cell Sciences Imaging Facility for providing fluorescence microscopy access and training; Stephanie Culler, Rachel Green,

III-61

Louis Maher III, and Melina Mathur for providing plasmid constructs; and Ryan Bloom and Melina Mathur for providing stable cell lines with integrated constructs. This work was supported by the National Institutes of Health (grant to Christina Smolke, RC1GM091298) and the Defense Advanced Research Projects Agency (grant to Christina Smolke, HR0011-11-2-0002).

References

1. Chang, A. L. et al. Synthetic RNA switches as a tool for temporal and spatial control over gene expression. *Curr. Opin. Biotechnol.* **23**, 679–688 (2012).
2. Nie, M. & Htun, H. Different modes and potencies of translational repression by sequence-specific RNA-protein interaction at the 5'-UTR. *Nucleic Acids Res.* **34**, 5528–5540 (2006).
3. Ausländer, S. et al. Programmable single-cell mammalian biocomputers. *Nature* **487**, 123–127 (2012).
4. Saito, H. et al. Synthetic translational regulation by an L7Ae-kink-turn RNP switch. *Nat. Chem. Biol.* **6**, 71–78 (2010).
5. Stapleton, J. a. et al. Feedback Control of Protein Expression in Mammalian Cells by Tunable Synthetic Translational Inhibition. *ACS Synth. Biol.* **1**, 83–88 (2012).
6. Endo, K. et al. Quantitative and simultaneous translational control of distinct mammalian mRNAs. *Nucleic Acids Res.* **41**, e135 (2013).
7. Liu, Y. et al. Synthesizing oncogenic signal-processing systems that function as both “signal counters” and “signal blockers” in cancer cells. *Mol. Biosyst.* **9**, 1909–1918 (2013).
8. Endo, K. et al. A versatile cis-acting inverter module for synthetic translational switches. *Nat. Commun.* **4**, 2393 (2013).
9. Saito, H. et al. Synthetic human cell fate regulation by protein-driven RNA switches. *Nat. Commun.* **2**, 160 (2011).
10. Culler, S. J. et al. Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins. *Science.* **330**, 1251–1255 (2010).
11. Win, M. N. & Smolke, C. D. A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14283–14288 (2007).
12. Michener, J. K. & Smolke, C. D. High-throughput enzyme evolution in *Saccharomyces cerevisiae* using a synthetic RNA switch. *Metab. Eng.* **14**, 306–316 (2012).
13. Galloway, K. E. et al. Dynamically reshaping signaling networks to program cell fate via genetic controllers. *Science.* **341**, 1235005 (2013).

14. Ausländer, S. et al. A ligand-dependent hammerhead ribozyme switch for controlling mammalian gene expression. *Mol. Biosyst.* **6**, 807–814 (2010).
15. Chen, Y. Y. et al. Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 8531–8536 (2010).
16. Wieland, M. et al. Engineering of ribozyme-based riboswitches for mammalian cells. *Methods* **56**, 351–357 (2012).
17. Wei, K. Y. et al. A yeast-based rapid prototype platform for gene control elements in mammalian cells. *Biotechnol. Bioeng.* **110**, 1201–1210 (2013).
18. Stripecke, R. et al. Proteins binding to 5' untranslated region sites: a general mechanism for translational regulation of mRNAs in human and yeast cells. *Mol. Cell. Biol.* **14**, 5898–5909 (1994).
19. Paraskeva, E. et al. A translational repression assay procedure (TRAP) for RNA-protein interactions in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 951–956 (1998).
20. Fusco, D. et al. Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr. Biol.* **13**, 161–167 (2003).
21. Hobson, D. & Uhlenbeck, O. C. Alanine scanning of MS2 coat protein reveals protein-phosphate contacts involved in thermodynamic hot spots. *J. Mol. Biol.* **356**, 613–624 (2006).
22. Keryer-Bibens, C. et al. Tethering of proteins to RNAs by bacteriophage proteins. *Biol. Cell* **100**, 125–138 (2008).
23. Khvorova, A. et al. Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. *Nat. Struct. Biol.* **10**, 708–712 (2003).
24. Li, X. et al. Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter. *J. Biol. Chem.* **273**, 34970–34975 (1998).
25. Beckett, D. & Uhlenbeck, O. C. Ribonucleoprotein complexes of R17 coat protein and a translational operator analog. *J. Mol. Biol.* **204**, 927–938 (1988).
26. Sugiyama, T. et al. Ribonucleoprotein Complexes formed between Bacteriophage MS2 RNA and MS2 Protein in vitro. *J. Mol. Biol.* **25**, 455–463 (1967).
27. LeCuyer, K. A. et al. Mutants of the bacteriophage MS2 coat protein that alter its cooperative binding to RNA. *Biochemistry* **34**, 10600–10606 (1995).
28. Peabody, D. S. & Lim, F. Complementation of RNA binding site mutations in MS2 coat protein heterodimers. *Nucleic Acids Res.* **24**, 2352–2359 (1996).

29. Wang, R. & Brattain, M. G. The maximal size of protein to diffuse through the nuclear pore is larger than 60kDa. *FEBS Lett.* **581**, 3164–3170 (2007).
30. Kalderon, D. et al. A short amino acid sequence able to specify nuclear location. *Cell* **39**, 499–509 (1984).
31. Wen, W. et al. Identification of a signal for rapid export of proteins from the nucleus. *Cell* **82**, 463–473 (1995).
32. Lim, F. et al. Translational repression and specific RNA binding by the coat protein of the *Pseudomonas* phage PP7. *J. Biol. Chem.* **276**, 22507–22513 (2001).
33. Tan, R. & Frankel, A. D. Structural variety of arginine-rich RNA-binding peptides. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 5282–5286 (1995).
34. Karin, M. et al. NF- κ B in cancer: from innocent bystander to major culprit. *Nat. Rev. Cancer* **2**, 301–310 (2002).
35. Lessard, L. et al. NF- κ B nuclear localization and its prognostic significance in prostate cancer. *BJU Int.* **91**, 417–420 (2003).
36. Kontos, S. et al. Inverse expression of estrogen receptor-beta and nuclear factor- κ B in urinary bladder carcinogenesis. *Int. J. Urol.* **17**, 801–809 (2010).
37. Koga, F. et al. ErbB2 and NF κ B overexpression as predictors of chemoradiation resistance and putative targets to overcome resistance in muscle-invasive bladder cancer. *PLoS One* **6**, e27616 (2011).
38. Kau, T. R. et al. Nuclear transport and cancer: from mechanism to intervention. *Nat. Rev. Cancer* **4**, 106–117 (2004).
39. Lebruska, L. L. & Maher, L. J. Selection and characterization of an RNA decoy for transcription factor NF- κ B. *Biochemistry* **38**, 3168–3174 (1999).
40. Wurster, S. & Maher, L. Selection and characterization of anti-NF- κ B p50 RNA aptamers. *RNA* **14**, 1037–1047 (2008).
41. Lee, H. K. et al. Modulation of oncogenic transcription and alternative splicing by beta-catenin and an RNA aptamer in colon cancer cells. *Cancer Res.* **66**, 10560–10566 (2006).
42. Choi, Y. S. et al. The RNA aptamer disrupts protein-protein interaction between beta-catenin and nuclear factor- κ B p50 and regulates the expression of C-reactive protein. *FEBS Lett.* **583**, 1415–1421 (2009).

43. Hogg, J. & Collins, K. RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation. *RNA* **13**, 868–880 (2007).
44. Ganchi, P. A. et al. I κ B/MAD-3 Masks the Nuclear Localization Signal of NF- κ B p65 and Requires the Transactivation Domain to Inhibit NF- κ B p65 DNA Binding. *Mol. Biol. Cell* **3**, 1339–1352 (1992).
45. Huxford, T. et al. The Crystal Structure of the I κ B α /NF- κ B Complex Reveals Mechanisms of NF- κ B Inactivation. *Cell* **95**, 759–770 (1998).
46. Huang, D.-B. et al. Crystal structure of NF- κ B (p50)₂ complexed to a high-affinity RNA aptamer. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9268–9273 (2003).
47. Ghosh, G. et al. NF- κ B regulation: lessons from structures. *Immunol. Rev.* **246**, 36–58 (2012).
48. Lee, H. K. & Jeong, S. β -Catenin stabilizes cyclooxygenase-2 mRNA by interacting with AU-rich elements of 3'-UTR. *Nucleic Acids Res.* **34**, 5705–5714 (2006).
49. Kim, I. et al. β -Catenin recognizes a specific RNA motif in the cyclooxygenase-2 mRNA 3'-UTR and interacts with HuR in colon cancer cells. *Nucleic Acids Res.* **40**, 6863–6872 (2012).
50. Liang, J. C. et al. A high-throughput, quantitative cell-based screen for efficient tailoring of RNA device activity. *Nucleic Acids Res.* **40**, e154 (2012).
51. Daugherty, P. S. et al. Flow cytometric screening of cell-based libraries. *J. Immunol. Methods* **243**, 211–227 (2000).
52. Kosuri, S. et al. Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 14024–14029 (2013).
53. Beisel, C. L. et al. Model-guided design of ligand-regulated RNAi for programmable control of gene expression. *Mol. Syst. Biol.* **4**, (2008).
54. Beisel, C. L. et al. Design of small molecule-responsive microRNAs based on structural requirements for Droscha processing. *Nucleic Acids Res.* **39**, 2981–2994 (2011).
55. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).

56. Mathews, D. H. Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *RNA* **10**, 1178–1190 (2004).

Supplementary Tables

Supplementary Table 3.1. DNA sequences of RNA devices.

RNA device	ligand	DNA sequence
sTRSV	N/A	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTGAGGACGAAACAGC
sTRSVctrl	N/A	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTGAGGACAAAACAGC
MS2-A1	MS2 coat protein	GCTGTCACCGGACTACACCATCAGGGTAGTGTGC TTTCCGGTCTGATGAGTCCGTGAGGACGAAACAG C
MS2-A2	MS2	GCTGTCACCGGATGTGCGTACACCATCAGGGTAC TTTCCGGTCTGATGAGTCCGTGAGGACGAAACAG C
MS2-A3	MS2	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTCGTACACCATCAGGGTACGGAGGACGAAA CAGC
MS2-A4	MS2	GCTGTCACCGGATGTGGTTTCCGGTCTGATGAGT CCGTCGTACACCATCAGGGTACGGAGGACGAAA CAGC
MS2-A5	MS2	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTCGTACACCATCAGGGTACGGAGGACGAAA CAGC
MS2-A6	MS2	GCTGTCACCGGATTCGGGATCCGGTCTGATGAGT CCGTCGTACACCATCAGGGTACGGAGGACGAAA CAGC
MS2-A7	MS2	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTGCGTACACCATCAGGGTACGAGGACGAAA CAGC
MS2-B1	MS2	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTTGTCCAGGATCACCGGACGGGACGGAGGA CGAAACAGC
MS2-B2	MS2	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTGTCCACCATCAGGGGACGGGACGGAGGA CGAAACAGC
MS2-B2ctrl	MS2	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTGTCCACCATCAGGGGACGGGACGGAGGA CAAAACAGC
MS2-B3	MS2	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTGGTCCACCATCAGGGGACTGGACTGAGGAC GAAACAGC
MS2-B4	MS2	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTCGTCCAGGATCACCGGACGGGACGGAGGA CGAAACAGC

MS2-B5	MS2	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTCGTCCTAGGATCACCAGGACGGGACGGAG GACGAAACAGC
MS2-B6	MS2	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTCGTCCTAGGATCACCAGGAAGGGACGGAG GACGAAACAGC
MS2-B7	MS2	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTTGTCCTAGGATCACCAGGAAGGGACGGAG GACGAAACAGC
MS2-B8	MS2	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTTGCGTAGGATCACCACGTGGCGCGGAGGAC GAAACAGC
MS2-B9	MS2	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTTGTAGGATCACCACACGGAGGACGAAACA GC
MS2-B10	MS2	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTGGTCCACCATCAGGGGACTGGACTGAGGAC GAAACAGC
MS2-B11	MS2	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTGTCCTAGGATCACCAGGAAGGGACGGAG GACGAAACAGC
MS2-C1	MS2	GCTGTCACCGGATGTGCTGCAGGATCACCGCATT TCCGGTCTGATGAGTCCGTGAGGACGAAACAGC
MS2-C1ctrl	MS2	GCTGTCACCGGATGTGCTGCAGGATCACCGCATT TCCGGTCTGATGAGTCCGTGAGGACAAAACAGC
MS2-C2	MS2	GCTGTCACCGGATGTGGTTTCCGGTCTGATGAGT CCGACCATCAGGAGGACGAAACAGC
MS2-C3	MS2	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGACCATCAGGAGGACGAAACAGC
MS2-C3ctrl	MS2	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGACCATCAGGAGGACAAAACAGC
MS2-C4	MS2	GCTGTCACCGGATTCGGGATCCGGTCTGATGAGT CCGACCATCAGGAGGACGAAACAGC
MS2-D1	MS2	TGCTGTACGATCACGACAGCGGGCTAAAGCCCGC TGTCACCGGATGTGCTTTCCGGTCTGATGAGTCC GTGAGGACGAAACAGCGGGCC
MS2-D2	MS2	TGCTGCACGATCACGGCAGCGAGCTAAAGCTCGC TGTCACCGGATGTGCTTTCCGGTCTGATGAGTCC GTGAGGACGAAACAGCGGGCC
MS2-D3	MS2	TGCTGCAGGATCACCGCAGCGAGCTAAAGCTCGC TGTCACCGGATGTGCTTTCCGGTCTGATGAGTCC GTGAGGACGAAACAGCGAC
MS2-D4	MS2	TGCTTCATGATCACAGGAGCGGCACTAAAGTGCC GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTGAGGACGAAACAGCGGGCC

MS2-D5	MS2	TAAAATAGTCATGATCACAGGCTGTCACCGGATG TGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAA CAGCCC
MS2-D6	MS2	TAAAATAGTTAGGATCACCGGCTGTCACCGGATG TGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAA CAGCCC
PP7-1	PP7 coat protein	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGAGTTTATATGGAAACAGGACGAAACAGC
PP7-2	PP7	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTGTTATATGGAAACGGGACGAAACAGC
PP7-3	PP7	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTGGTATATGGACCGGGACGAAACAGC
Lambda-1	Lambda N 1-22 peptide	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTGTCGAAGAGGACGGGACGGAGGACGAA ACAGC
Lambda-2	Lambda N	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTGTCCTGAAGAAGGAAGGGACGGAGGACG AACAGC
Lambda-3	Lambda N	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTAGAGGACGAAACAGC
Lambda-4	Lambda N	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCTGTAGAAGGACGAAACAGC
p50-1	NF- κ Bp50	GCTGTCACCGGATGTGCTTCCGGTCTGATGAGT CCGTTGGTATCCTGAAACTGTTATAAGGTTGGCC GATGTTGGAGGACGAAACAGC
p50-2	p50	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTGGTATCCTGAAACTGTTATAAGGTTGGCC GATGTTGGAGGACGAAACAGC
p50-3	p50	GCTGTCACCGGATGTGCTTCCGGTCTGATGAGT CCGTGGGTATCCTGAAACTGTTATAAGGTTGGCC GATGTTGGAGGACGAAACAGC
p50-4	p50	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTGGGTATCCTGAAACTGTTATAAGGTTGGCC GATGTTGGAGGACGAAACAGC
p50-5	p50	GCTGTCACCGGATGTGCTTCCGGTCTGATGAGT CCGTTGGTATCCTGAAACTGTTATAAGGTTGGCC GATGTGTGGAGGACGAAACAGC
p50-6	p50	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTGGTATCCTGAAACTGTTATAAGGTTGGCC GATGTGTGGAGGACGAAACAGC
p50-7	p50	GCTGTCACCGGATGTGCTTCCGGTCTGATGAGT CCGTTGTCCATCCTGAAACTGTTATAAGGTTGGC CGATGGACGGATGGACGGAGGACGAAACAGC
p50-8	p50	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTGTCCATCCTGAAACTGTTATAAGGTTGGC CGATGGACGGATGGACGGAGGACGAAACAGC

III-70

p50-9	p50	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTTGTCCATCCTGAAACTGTTATAAGGTTGGC CGATGGACAGATGGACGGAGGACGAAACAGC
p50-10	p50	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTGTCCATCCTGAAACTGTTATAAGGTTGGC CGATGGACAGATGGACGGAGGACGAAACAGC
p50-11	p50	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTCATCCTGAAACTGTTATAAGGTTGGCCGAT GGGACGGAGGACGAAACAGC
p50-12	p50	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTGTCCTTGAAACTGTAAGGTTGGCGGACGGG ACGAGGACGAAACAGC
p50-13	p50	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTCATCCTGAAACTGTTATAAGGTTGGCCGAT GGGACGAGGACGAAACAGC
p65-1	NF- κ Bp65	GCTGTCACCGGATGTGCTCGATGAGTCCGCGAGG TGCCGAACCTCCATTGGGGTTCGGTTTCCGGTCTG ATGAGTCCGTGAGGACGAAACAGC
p65-2	p65	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTTTGATTCGATGAGTCCGCGAGGTGCCGAAC CTCCATTGGGGTTCGAGAGGACGAAACAGC
p65-3	p65	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTTGATTCGATGAGTCCGCGAGGTGCCGAAC CTCCATTGGGGTTCGAGAGGACGAAACAGC
p65-4	p65	GCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTTTGATCGATGAGTCCGCGAGGTGCCGAACC TCCATTGGGGTTCGAGAGGACGAAACAGC
p65-5	p65	GCTGTCACCGGAATCAAGGTCCGGTCTGATGAGT CCGTTTGATCGATGAGTCCGCGAGGTGCCGAACC TCCATTGGGGTTCGAGAGGACGAAACAGC
Beat-A1	β -catenin	GACGTATGAGACTATGGACGCTATAGGCACACCG GATACTTTAACGTCTCACTGATGAGGCCATGGCA GGCCGAAACGTC
Beat-A2	β -catenin	TACGTCTGAGCGTATGGACGCTATAGGCACACCG GATACTTTAACCGCTCACTGAAGATGGCCCGGTA GGGCCGAAACGTA
Beat-A3	β -catenin	AAGAGGTCGGCACCTATGGACGCTATAGGCACA CCGGATACTTTAACGGTGTCTGATGAAGATCCA TGACAGGATCGAAACCTCTT
Beat-A4	β -catenin	CGCTGTCTGTA CT TATGGACGCTATAGGCACACC GGATACTTTAACAGTACACTGACGAGTCCCTAAA GGACGAAACAGCG
Beat-B1	β -catenin	GGCCGATCTATGGACGCTATAGGCACACCGGATA CTTTAACGATTGGCTATAAAAGCTGTCACCGGAT GTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAA ACAGCC

Bcat-B2	β -catenin	GCCGATCTATGGACGCTATAGGCACACCCGGATAC TTTAACGATTGGCATAAAAGCTGTCACCCGGATGT GCTTTCCGGTCTGATGAGTCCGTGAGGACGAAAC AGCC
Bcat-B3	β -catenin	TAGGCCGATCTATGGACGCTATAGGCACACCCGGA TACTTTAACGATTGGCTAAAAGCTGTCACCCGGAT GTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAA ACAGCC
Bcat-B4	β -catenin	TAAGCCGATCTATGGACGCTATAGGCACACCCGGA TACTTTAACGATTGGCAAAGCTGTCACCCGGATGT GCTTTCCGGTCTGATGAGTCCGTGAGGACGAAAC AGCC
Bcat-B5	β -catenin	TAAAAACCAGCATCTATGGACGCTATAGGCACAC CGGATACTTTAACGATGCTGTCACCCGGATGTGCT TTCCGGTCTGATGAGTCCGTGAGGACGAAACAGC ATC
Bcat-B6	β -catenin	TAAAATCGCCGATCTATGGACGCTATAGGCACAC CGGATACTTTAACGATTGGCTACCCGGATGTGCT TTCCGGTCTGATGAGTCCGTGAGGACGAAAGCCA

Supplementary Table 3.2. Free energies (ΔG , kcal/mol) of individual conformations (ribozyme cleavage-active and -inactive) and the energy difference ($\Delta\Delta G$, kcal/mol) predicted by RNAstructure 5.3⁵⁶.

RNA device	aptamer unformed		aptamer formed		inactive - active
	active	inactive	inactive	active	
sTRSV	-19.3		N/A		N/A
sTRSVctrl		-19.3	N/A		N/A
MS2-A1	N/A			-23.2	N/A
MS2-A2	N/A			-24.5	N/A
MS2-A3	N/A			-31.1	N/A
MS2-A4	N/A			-31.1	N/A
MS2-A5	N/A			-30.2	N/A
MS2-A6	N/A			-32.4	N/A
MS2-A7	N/A			-27.1	N/A
MS2-B1	-27.7		-26.9		0.8
MS2-B2	-28.7		-27.3		1.4
MS2-B2ctrl		-28.7	-27.3		N/A
MS2-B3	-30.6		-29.1		1.5
MS2-B4	-29.7		-28.5		1.2

III-72

MS2-B5	-30.4		-30.3		0.1
MS2-B6	-30.4		-27.9		2.5
MS2-B7	-28.5		-26.1		2.4
MS2-B8	-27.4		-25.5		1.9
MS2-B9	-22.7		-21.6		1.1
MS2-B10	-29.7		-28.2		1.5
MS2-B11	-27.5		-25.2		2.3
MS2-C1		-24.6		-22.3	-2.3
MS2-C1ctrl		-24.6	-22.3		N/A
MS2-C2	-20.6		N/A		N/A
MS2-C3	-19.7		N/A		N/A
MS2-C3ctrl		-19.7	N/A		N/A
MS2-C4	-21.9		N/A		N/A
MS2-D1		-41.8		-40.7	-1.1
MS2-D2		-39.1		-39.8	0.7
MS2-D3		-36.2		-35.6	-0.6
MS2-D4		-34.7		-34.3	-0.4
MS2-D5	-23.6		-23.0		0.6
MS2-D6	-23.8		-22.5		1.3
PP7-1	-19.7		N/A		N/A
PP7-2	-19.6		N/A		N/A
PP7-3	-22.5		N/A		N/A
Lambda-1	-27.6		-25.2		2.4
Lambda-2	-28.2		-23.9		4.3
Lambda-3	-18.3		N/A		N/A
Lambda-4	-19.6		N/A		N/A
p50-1		-25.3		-24.8	-0.5
p50-2		-24.4		-23.9	-0.5
p50-3		-24.0		-23.7	-0.3
p50-4		-23.7		-23.4	-0.3
p50-5		-25.1		-23.0	-2.1
p50-6		-24.8		-22.7	-2.1
p50-7	-38.0		-36.3		1.7
p50-8	-37.7		-36.0		1.7
p50-9	-36.8		-36.2		0.6
p50-10	-36.5		-35.9		0.6
p50-11	-27.8		-26.1		1.7
p50-12	-30.5		-28.3		2.2
p50-13	-26.8		-26.1		0.7
p65-1		-35.0		-34.4	-0.6
p65-2	-36.0		-33.7		2.3

III-73

p65-3	-35.1		-32.8		2.3
p65-4	-34.3		-34.1		0.2
p65-5	-33.4		-33.2		0.2
Bcat-A1	-21.0		N/A		N/A
Bcat-A2	-23.5		N/A		N/A
Bcat-A3	-23.8		N/A		N/A
Bcat-A4	-18.2		N/A		N/A
Bcat-B1		-31.3		-31.9	0.6
Bcat-B2		-30.3		-31.3	1.0
Bcat-B3		-33.1		-32.2	-0.9
Bcat-B4		-30.9		-31.3	0.4
Bcat-B5	-29.7		-28.8		0.9
Bcat-B6	-27.2		-25.9		1.3

Supplementary Table 3.3. Primer and oligonucleotide sequences.

Name	DNA Sequence
5' spacer	AATAAATAAAA
3' spacer	CAAATAAACAAACACTC
S 5' spacer	AATAAA
S 3' spacer	AAATAAACAAACACTC
lambda N	TCGAGATGGACGCCAGACCAGAAGGAGAGAGAGGAG AGCCGAGAAGCAGGCCAGTGGAAGGCCGCAACTAG CGGC
insertNA	CTGGCTAAAGGTGCGT
No NLS A/X Fwd	CAACTCGAGATGGTGGCTTCTAACTTTACTCAG
DsRed A/X Rev	CAAGGGCCCGCCGCTACTGGGAGC
MS2 A/X Rev	CAAGGGCCCGCCGCTAGTAGATGCCG
EF1BFP Fwd	CAAAGATCTGGATCTGCGATCGC
EF1BFP Rev	CAACCTAGGTCAATTAAGCTTGTGCCCCAG
MS2 NotI F	CAAGCGGCCGCTCGAGATGGCTTCTAACTTTACTCAGTT CGTTCTC
MS2 ApaI R	CAAGGGCCCGCCGCTAGTAGATGCCGGAGTTTGCT
NLS MS2 F	CAAGCGGCCGCTCGAGATGCCAAAGAAGAAGCGCAAA GTGGCTTCTAACTTTACTCAGTTCGTTCTC
2MS2mut R	AGCGGGTTTAAACGGGCCCGCCGCTA
2MS2mut F	GCAGATATCCAGCACAGTGGCGGCCGCTCGAGATG

NES MS2 R	CAAGGGCCCGCCGCTAGATGTCCAGTCCGGCCAGCTTC AGGGCCAGCTCGTTGTAGATGCCGGAGTTTGCT
DsRed GF	CCCGATTCCCTCGGCAATCGCAGCAAACCTCCGGCATCT ACG
DsRed NES R	AACTAGAAGGCACAGTCGAGGCTGATCAGCGGGTTTAA ACGGGCCCCGCCGCTAGATGTCCAGTCCGGCCAGCTTCA GGGCCAGCTCGTTCTGGGAGCCGGAGTG
DsRed GR	AACTAGAAGGCACAGTCGAGG
GF Clover	CCTCGAGTATTCGCCACCATGGTGAGCAAGGGCG
GR Clover	GCGCGCCTTACCTAGGTTACTTGTACAGCTCGTCCATGC
GF EF1	GCATGGACGAGCTGTACAAGTAACCTAGGTAAGGGCGC C
GR EF1	CGCCCTTGCTCACCATGGTGGCGAATACTCGAGG
PP7 NotI F	CAAGCGGCCGCTCGAGATGGCCAAAACCATCGTTCT
PP7 ApaI R	CAAGGGCCCGCCGCTAGGAACGGCCCAGCG
p50 NotI F	CAAGCGGCCGCTCGAGATGGCAGAAGATGATCCATATT TGGGAAG
p50 ApaI R	CAAGGGCCCGCCGCTAGTCATCACTTTTGTACAAACCTT CA
p65 NotI F	CAAGCGGCCGCTCGAGATGGACGATCTGTTTCCCCT
p65 ApaI R	CAAGGGCCCGCCGCTAGGTCCTTTTCGCCTTCTCTTC
b-cat NotI F	CAAGCGGCCGCTCGAGATGCGTGCAATCCCTGAACTGA
b-cat ApaI R	CAAGGGCCCGCCGCTACTTGTCTCAGACATTCGGAAC

Supplementary Table 3.4. Plasmid constructs.

Plasmid	Description
pCS2304	pcDNA5/FRT with d2EGFP, bGHpA, and CMV-TetO2 inserted between AflII/KpnI
pCS2359	pCS2304 with MS2-DsRedMonomer fusion protein inserted between XhoI/ApaI
pCS2394	pCS2359 with sTRSV with spacers inserted between AvrII/AscI
pCS2380	pCS2359 with MS2-A1 with spacers inserted between AvrII/AscI
pCS2381	pCS2359 with MS2-A2 with spacers inserted between AvrII/AscI
pCS2382	pCS2359 with MS2-A5 with spacers inserted between AvrII/AscI

III-75

pCS2383	pCS2359 with MS2-A6 with spacers inserted between AvrII/AscI
pCS2406	pCS2359 with d2EGFP removed from NheI/AvrII
pCS2409	pCS2406 with MS2 inserted between XhoI/ApaI
pCS1	blank plasmid with no mammalian promoters (Maung Win)
pCS339	pcDNA3.1(+) with chloramphenicol acetyltransferase (Chase Beisel)
pCS1392	pcDNA5/FRT with FLAG-NLS-MS2-DsRedMonomer (Stephanie Culler)
pCS2585	pcDNA5/FRT with EF1 α -BFP-HSVTK (Melina Mathur)
pCS2595	pCS2304 with EF1 α -BFP inserted between BglII/AvrII and MS2 inserted between XhoI/ApaI
pCS2602	pCS2595 with sTRSVctrl with spacers inserted between AvrII/AscI
pCS2601	pCS2595 with sTRSV with spacers inserted between AvrII/AscI
pCS2615	pCS2595 with MS2-B1 with spacers inserted between AvrII/AscI
pCS2616	pCS2595 with MS2-B2 with spacers inserted between AvrII/AscI
pCS1697	pCS2595 with MS2-B2ctrl with spacers inserted between AvrII/AscI
pCS2621	pCS2595 with MS2-B3 with spacers inserted between AvrII/AscI
pCS2610	pCS2595 with MS2-B4 with spacers inserted between AvrII/AscI
pCS2611	pCS2595 with MS2-B6 with spacers inserted between AvrII/AscI
pCS2612	pCS2595 with MS2-B7 with spacers inserted between AvrII/AscI
pCS2613	pCS2595 with MS2-B8 with spacers inserted between AvrII/AscI
pCS2614	pCS2595 with MS2-B9 with spacers inserted between AvrII/AscI
pCS2617	pCS2595 with MS2-C1 with spacers inserted between AvrII/AscI
pCS1698	pCS2595 with MS2-C1ctrl with spacers inserted between AvrII/AscI
pCS2618	pCS2595 with MS2-C2 with spacers inserted between AvrII/AscI
pCS2619	pCS2595 with MS2-C3 with spacers inserted between AvrII/AscI
pCS2385	pCS2595 with MS2-C3ctrl with spacers inserted between AvrII/AscI

pCS2620	pCS2595 with MS2-C4 with spacers inserted between AvrII/AscI
pCS2595+MS2-D1	pCS2595 with MS2-D1 with S spacers inserted between AvrII/AscI
pCS2595+MS2-D2	pCS2595 with MS2-D2 with S spacers inserted between AvrII/AscI
pCS2606	pCS2595 with MS2-D3 with S spacers inserted between AvrII/AscI
pCS2607	pCS2595 with MS2-D4 with S spacers inserted between AvrII/AscI
pCS2608	pCS2595 with MS2-D5 with S spacers inserted between AvrII/AscI
pCS2609	pCS2595 with MS2-D6 with S spacers inserted between AvrII/AscI
pCS2631	pCS2595 with MS2mut inserted between NotI/ApaI
pCS2632	pCS2631 with MS2-B1 with spacers inserted between AvrII/AscI
pCS2633	pCS2631 with MS2-B2 with spacers inserted between AvrII/AscI
pCS2650	pCS2631 with MS2-B3 with spacers inserted between AvrII/AscI
pCS2651	pCS2631 with MS2-B7 with spacers inserted between AvrII/AscI
pCS2634	pCS2631 with MS2-C1 with spacers inserted between AvrII/AscI
pCS2635	pCS2631 with MS2-C3 with spacers inserted between AvrII/AscI
pCS2653	pCS2631 with MS2-D5 with S spacers inserted between AvrII/AscI
pCS2686	pCS2595 with 2MS2mut inserted between NotI/ApaI
pCS2688	pCS2686 with sTRSVctrl with spacers inserted between AvrII/AscI
pCS2687	pCS2686 with sTRSV with spacers inserted between AvrII/AscI
pCS2690	pCS2686 with MS2-B1 with spacers inserted between AvrII/AscI
pCS2691	pCS2686 with MS2-B2 with spacers inserted between AvrII/AscI
pCS2694	pCS2686 with MS2-B3 with spacers inserted between AvrII/AscI
pCS2695	pCS2686 with MS2-B7 with spacers inserted between AvrII/AscI
pCS2698	pCS2686 with MS2-B10 with spacers inserted between AvrII/AscI

III-77

pCS2699	pCS2686 with MS2-B11 with spacers inserted between AvrII/AscI
pCS2692	pCS2686 with MS2-C1 with spacers inserted between AvrII/AscI
pCS2693	pCS2686 with MS2-C3 with spacers inserted between AvrII/AscI
pCS2697	pCS2686 with MS2-C4 with spacers inserted between AvrII/AscI
pCS2696	pCS2686 with MS2-D5 with S spacers inserted between AvrII/AscI
pCS2747	pCS2595 with NLS-2MS2mut inserted between NotI/ApaI
pCS2749	pCS2747 with sTRSVctrl with spacers inserted between AvrII/AscI
pCS2748	pCS2747 with sTRSV with spacers inserted between AvrII/AscI
pCS2750	pCS2747 with MS2-B1 with spacers inserted between AvrII/AscI
pCS2751	pCS2747 with MS2-B2 with spacers inserted between AvrII/AscI
pCS2758	pCS2747 with MS2-B11 with spacers inserted between AvrII/AscI
pCS2752	pCS2747 with MS2-C1 with spacers inserted between AvrII/AscI
pCS2753	pCS2747 with MS2-C3 with spacers inserted between AvrII/AscI
pCS2759	pCS2747 with MS2-D5 with S spacers inserted between AvrII/AscI
pCS2787	pCS2595 with 2MS2mut-NES inserted between NotI/ApaI
pCS2789	pCS2787 with sTRSVctrl with spacers inserted between AvrII/AscI
pCS2788	pCS2787 with sTRSV with spacers inserted between AvrII/AscI
pCS2790	pCS2787 with MS2-B1 with spacers inserted between AvrII/AscI
pCS2791	pCS2787 with MS2-B2 with spacers inserted between AvrII/AscI
pCS2798	pCS2787 with MS2-B11 with spacers inserted between AvrII/AscI
pCS2792	pCS2787 with MS2-C1 with spacers inserted between AvrII/AscI
pCS2793	pCS2787 with MS2-C3 with spacers inserted between AvrII/AscI
pCS2799	pCS2787 with MS2-D5 with S spacers inserted between AvrII/AscI
pCS2897	pCS2595 with 2MS2mut-DsRedMonomer

pCS2899	pCS2897 with sTRSVctrl with spacers inserted between AvrII/AscI
pCS2898	pCS2897 with sTRSV with spacers inserted between AvrII/AscI
pCS2901	pCS2897 with MS2-B11 with spacers inserted between AvrII/AscI
pCS2900	pCS2897 with MS2-C1 with spacers inserted between AvrII/AscI
pCS2902	pCS2595 with NLS-2MS2mut-DsRedMonomer
pCS2904	pCS2902 with sTRSVctrl with spacers inserted between AvrII/AscI
pCS2903	pCS2902 with sTRSV with spacers inserted between AvrII/AscI
pCS2906	pCS2902 with MS2-B11 with spacers inserted between AvrII/AscI
pCS2905	pCS2902 with MS2-C1 with spacers inserted between AvrII/AscI
pCS2907	pCS2595 with 2MS2mut-DsRedMonomer-NES
pCS2909	pCS2907 with sTRSVctrl with spacers inserted between AvrII/AscI
pCS2908	pCS2907 with sTRSV with spacers inserted between AvrII/AscI
pCS2911	pCS2907 with MS2-B11 with spacers inserted between AvrII/AscI
pCS2910	pCS2907 with MS2-C1 with spacers inserted between AvrII/AscI
pCS1576	pcDNA3.1(+) with DsRed-Express
pCS2391	pCS2595 with EF1 α -Clover inserted between BglII/AvrII
pCS2847	pCS2595 with PP7 inserted between NotI/ApaI
pCS2847+sTRSVctrl	pCS2847 with sTRSVctrl with spacers inserted between AvrII/AscI
pCS2847+sTRSV	pCS2847 with sTRSV with spacers inserted between AvrII/AscI
pCS2847+PP7-1	pCS2847 with PP7-1 with spacers inserted between AvrII/AscI
pCS2847+PP7-2	pCS2847 with PP7-2 with spacers inserted between AvrII/AscI
pCS2847+PP7-3	pCS2847 with PP7-3 with spacers inserted between AvrII/AscI
pCS2816	pCS2595 with lambda N inserted between NotI/ApaI
pCS2816+sTRSVctrl	pCS2816 with sTRSVctrl with spacers inserted between AvrII/AscI
pCS2816+sTRSV	pCS2816 with sTRSV with spacers inserted between AvrII/AscI
pCS2816+Lambda-1	pCS2816 with Lambda-1 with spacers inserted between AvrII/AscI

pCS2816+Lambda-2	pCS2816 with Lambda-2 with spacers inserted between AvrII/AscI
pCS2816+Lambda-3	pCS2816 with Lambda-3 with spacers inserted between AvrII/AscI
pCS2816+Lambda-4	pCS2816 with Lambda-4 with spacers inserted between AvrII/AscI
pCS2397	pCS2595 with insertNA inserted between NotI/ApaI
pCS2666	pCS2397 with sTRSVctrl with spacers inserted between AvrII/AscI
pCS2665	pCS2397 with sTRSV with spacers inserted between AvrII/AscI
pCS2672	pCS2397 with p50-1 with spacers inserted between AvrII/AscI
pCS2673	pCS2397 with p50-2 with spacers inserted between AvrII/AscI
pCS2674	pCS2397 with p50-3 with spacers inserted between AvrII/AscI
pCS2675	pCS2397 with p50-4 with spacers inserted between AvrII/AscI
pCS2676	pCS2397 with p50-5 with spacers inserted between AvrII/AscI
pCS2677	pCS2397 with p50-6 with spacers inserted between AvrII/AscI
pCS2668	pCS2397 with p50-7 with spacers inserted between AvrII/AscI
pCS2669	pCS2397 with p50-8 with spacers inserted between AvrII/AscI
pCS2670	pCS2397 with p50-9 with spacers inserted between AvrII/AscI
pCS2671	pCS2397 with p50-10 with spacers inserted between AvrII/AscI
pCS2683	pCS2397 with p50-11 with spacers inserted between AvrII/AscI
pCS2684	pCS2397 with p50-12 with spacers inserted between AvrII/AscI
pCS2685	pCS2397 with p50-13 with spacers inserted between AvrII/AscI
pCS2682	pCS2397 with p65-1 with spacers inserted between AvrII/AscI
pCS2678	pCS2397 with p65-2 with spacers inserted between AvrII/AscI
pCS2679	pCS2397 with p65-3 with spacers inserted between AvrII/AscI
pCS2680	pCS2397 with p65-4 with spacers inserted between AvrII/AscI
pCS2681	pCS2397 with p65-5 with spacers inserted between AvrII/AscI
pCS2604	pCS2391 with p50 inserted between NotI/ApaI
pCS2605	pCS2391 with p65 inserted between NotI/ApaI
pCS2766	pCS2397 with Bcat-A1 with spacers inserted between AvrII/AscI
pCS2767	pCS2397 with Bcat-A2 with spacers inserted between AvrII/AscI
pCS2768	pCS2397 with Bcat-A3 with spacers inserted between AvrII/AscI
pCS2769	pCS2397 with Bcat-A4 with spacers inserted between AvrII/AscI

III-80

pCS2397+Bcat-B1	pCS2397 with Bcat-B1 with S spacers inserted between AvrII/AscI
pCS2397+Bcat-B2	pCS2397 with Bcat-B2 with S spacers inserted between AvrII/AscI
pCS2397+Bcat-B3	pCS2397 with Bcat-B3 with S spacers inserted between AvrII/AscI
pCS2397+Bcat-B4	pCS2397 with Bcat-B4 with S spacers inserted between AvrII/AscI
pCS2397+Bcat-B5	pCS2397 with Bcat-B5 with S spacers inserted between AvrII/AscI
pCS2397+Bcat-B6	pCS2397 with Bcat-B6 with S spacers inserted between AvrII/AscI
pCS2824	pCS2595 with β -catenin Arm 1-12 inserted between NotI/ApaI
pCS2850	pCS2824 with sTRSVctrl with spacers inserted between AvrII/AscI
pCS2849	pCS2824 with sTRSV with spacers inserted between AvrII/AscI
pCS2824+Bcat-A1	pCS2824 with Bcat-A1 with spacers inserted between AvrII/AscI
pCS2824+Bcat-A2	pCS2824 with Bcat-A2 with spacers inserted between AvrII/AscI
pCS2824+Bcat-A3	pCS2824 with Bcat-A3 with spacers inserted between AvrII/AscI
pCS2848	pCS2824 with Bcat-A4 with spacers inserted between AvrII/AscI
pCS2824+Bcat-B1	pCS2824 with Bcat-B1 with S spacers inserted between AvrII/AscI
pCS2824+Bcat-B2	pCS2824 with Bcat-B2 with S spacers inserted between AvrII/AscI
pCS2824+Bcat-B3	pCS2824 with Bcat-B3 with S spacers inserted between AvrII/AscI
pCS2824+Bcat-B4	pCS2824 with Bcat-B4 with S spacers inserted between AvrII/AscI
pCS2824+Bcat-B5	pCS2824 with Bcat-B5 with S spacers inserted between AvrII/AscI
pCS2824+Bcat-B6	pCS2824 with Bcat-B6 with S spacers inserted between AvrII/AscI

Supplementary Table 3.5. Human cell lines with stably integrated constructs.

Parental line	Integrated plasmid construct
Flp-In T-REx HEK293	pCS2359
Flp-In T-REx HEK293	pCS2380
Flp-In T-REx HEK293	pCS2381
Flp-In T-REx HEK293	pCS2382
Flp-In T-REx HEK293	pCS2383
Flp-In T-REx HEK293	EF1 α -GFP (Ryan Bloom)
Flp-In HEK293	pCS2585 (Melina Mathur)
Flp-In T-REx HEK293	pCS2595
Flp-In T-REx HEK293	pCS2686
Flp-In T-REx HEK293	pCS2688
Flp-In T-REx HEK293	pCS2699
Flp-In T-REx HEK293	pCS2692
Flp-In T-REx HEK293	pCS2693
Flp-In T-REx HEK293	pCS2696
Flp-In T-REx HEK293	pCS2747
Flp-In T-REx HEK293	pCS2749
Flp-In T-REx HEK293	pCS2758
Flp-In T-REx HEK293	pCS2752
Flp-In T-REx HEK293	pCS2753
Flp-In T-REx HEK293	pCS2759
Flp-In T-REx HEK293	pCS2787
Flp-In T-REx HEK293	pCS2789
Flp-In T-REx HEK293	pCS2798
Flp-In T-REx HEK293	pCS2792
Flp-In T-REx HEK293	pCS2793
Flp-In T-REx HEK293	pCS2799
Flp-In T-REx HEK293	pCS2899
Flp-In T-REx HEK293	pCS2898
Flp-In T-REx HEK293	pCS2901
Flp-In T-REx HEK293	pCS2900
Flp-In T-REx HEK293	pCS2904
Flp-In T-REx HEK293	pCS2903
Flp-In T-REx HEK293	pCS2906
Flp-In T-REx HEK293	pCS2905
Flp-In T-REx HEK293	pCS2909
Flp-In T-REx HEK293	pCS2908
Flp-In T-REx HEK293	pCS2911

Flp-In T-REx HEK293	pCS2910
Flp-In T-REx HEK293	pCS2824
Flp-In T-REx HEK293	pCS2850
Flp-In T-REx HEK293	pCS2849
Flp-In T-REx HEK293	pCS2848

Supplementary Table 3.6. Alternate names of RNA devices.

RNA device	Name in notebooks	RNA device	Name in notebooks
MS2-A1	LI AU	PP7-3	PP7-3
MS2-A2	LI CU	Lambda-1	L2b8-a1-lambda
MS2-A3	LII UG	Lambda-2	L85-a1-lambda
MS2-A4	LII UG CtoG	Lambda-3	lambda-4
MS2-A5	LII UG a1	Lambda-4	lambda-5
MS2-A6	LII UG a14	p50-1	p50-OFF1
MS2-A7	LII GA	p50-2	p50-OFF1-a1
MS2-B1	D1	p50-3	p50-OFF2
MS2-B2	D2	p50-4	p50-OFF2-a1
MS2-B2ctrl	D2ctrl	p50-5	p50-OFF3
MS2-B3	D7	p50-6	p50-OFF3-a1
MS2-B4	L81	p50-7	p50-ON1
MS2-B5	L83	p50-8	p50-ON1-a1
MS2-B6	L84	p50-9	p50-ON2
MS2-B7	L85	p50-10	p50-ON2-a1
MS2-B8	L61	p50-11	U12N11
MS2-B9	L41	p50-12	U1N1
MS2-B10	D7-a1	p50-13	U11N11
MS2-B11	L85-a1	p65-1	p65-OFF1
MS2-C1	D3	p65-2	p65-ON1
MS2-C1ctrl	D3ctrl	p65-3	p65-ON1-a1
MS2-C2	D4	p65-4	p65-ON2
MS2-C3	D5	p65-5	p65-ON2-a1
MS2-C3ctrl	D5ctrl	Bcat-A1	Bcat-sLTSV-
MS2-C4	D6	Bcat-A2	Bcat-sLTSV+
MS2-D1	J1	Bcat-A3	Bcat-CChMvd-
MS2-D2	J2	Bcat-A4	Bcat-SCMoV+

III-83

MS2-D3	J3
MS2-D4	J4
MS2-D5	J5
MS2-D6	J6
PP7-1	D5-PP7
PP7-2	PP7-2

Bcat-B1	Bcat-OFF3
Bcat-B2	Bcat-OFF4
Bcat-B3	Bcat-OFF5
Bcat-B4	Bcat-OFF6
Bcat-B5	Bcat-ON3
Bcat-B6	Bcat-ON4

Chapter 4

Conclusions

IV-2

The emerging field of synthetic biology has produced a vast array of engineered molecular devices, enabling investigation of cellular function and programmed control of new phenotypic behaviors in biological systems¹⁻⁴. These devices are generally composed of protein or RNA, two biological macromolecules whose sequence determines their three-dimensional shape, dictating their ability to bind to other molecules and catalyze chemical reactions. RNA and protein engineering have greatly expanded the capabilities of these macromolecules, enabling functions not found in natural biological systems^{3,5-8}.

Synthetic molecular devices have been used to regulate gene expression in a wide variety of organisms, from prokaryotes to microbial eukaryotes to humans⁹⁻¹¹. Some of these genetic control platforms are able to process molecular input into increases or decreases in gene expression output by combining a sensor component with an actuator component^{1,12}. Such platforms exhibit the greatest utility when the components are easy to design and optimize, and when different components can be integrated together in predictable ways without disrupting their individual functions.

RNA is particularly well suited as a substrate for the implementation of molecular gene-regulatory devices. RNA molecules can hybridize with RNA and DNA through base-pairing interactions, and bind to small molecules and proteins by adopting specific conformations^{13,14}. They are also able to catalyze various chemical reactions, including the lysis of phosphodiester bonds^{15,16}. The binding and catalytic functions of RNA strands are largely dictated by their secondary structure, which can be predicted by computational models of RNA folding¹⁷⁻¹⁹. In contrast, protein function largely depends on complex tertiary interactions, which are currently far more challenging to predict from

the primary sequence alone. Furthermore, the ability of RNA to be replicated by reverse transcription and PCR enables the facile *in vitro* selection of RNA molecules with novel functions from large libraries of different sequences^{20,21}.

The ability of RNA enzymes to cleave phosphodiester bonds is exploited in the engineering of the ribozyme switch platform, in which cleavage of an mRNA strand by a hammerhead ribozyme causes silencing of the encoded gene in response to ligand binding to an aptamer¹². With the aid of structure prediction software, ribozyme switches were designed to adopt distinct cleavage-active and cleavage-inactive conformations, with ligand binding stabilizing the conformation in which the aptamer sensor component is properly formed¹². Both ON and OFF switches were demonstrated to regulate gene expression in yeast and mammalian cells and, importantly, replacement of the aptamer component to sense an alternate ligand did not require extensive redesign of the device^{12,22,23}. However, the platform was limited to the regulation of transgenes in response to small molecule inputs.

We attempted to extend the capabilities of the ribozyme switch platform to two new functions: the regulation of endogenous genes and the sensing of protein inputs. We were unable to demonstrate ribozyme activity *in trans*, and the limitations we discovered suggest that this platform is not as promising as other trans-acting platforms such as those based on RNAi²⁴⁻²⁶ and CRISPRi²⁷. We were successful, however, in developing novel protein-responsive ribozyme switches for regulating genes *in cis* in human cells. We demonstrated a higher level of ligand-responsiveness than previously described small-molecule-responsive ribozyme switches in mammalian systems, and we showed that cytoplasmic and nuclear localization of ligand were each sufficient to elicit switching

activity. We also demonstrated the versatility of our switch platform with a ribozyme switch responsive to an alternative protein ligand.

In our attempts to develop new devices responsive to various protein ligands, we found that integration of aptamers into the platform is a challenging process. Further study is needed in order to extend our platform to diverse ligands, making the process of generating new devices more reliable and straightforward. We rationally designed each device presented here, but in the future a wider sequence space could be explored using high-throughput *in vivo* screening methods^{20,28,29} to assay large libraries of randomized devices. As new sensor components are generated by *in vitro* selection, we hope that improved screening strategies will enable them to be integrated into our switch platform.

Our ribozyme switch is able to respond to proteins in either the nucleus or the cytoplasm, while previously described mammalian gene-regulatory devices have required specific localization of ligand in order to produce a switching response. Our platform is therefore unable to detect changes in protein distribution across subcellular compartments, but it is more versatile than previous platforms in that ligand input choice is not restricted to proteins localized to just one compartment.

We developed a device responsive to β -catenin, a signaling protein with an important role in cancer³⁰. This device and other ribozyme switches that respond to disease markers could be used to noninvasively detect diseased cellular states. Furthermore, such switches could be used to control cell fate by, for example, regulating the expression of a proapoptotic transgene. In this way a genetically encoded therapeutic effect could be targeted to diseased cells while leaving healthy cells unaffected. Coupling our device with other synthetic biology components such as positive feedback

or amplifier systems could expand the dynamic range of switch response and enable tuning of activity to match application-specific phenotypic thresholds. As the field of synthetic biology continues to advance, we hope the molecular device platform we have developed will be a useful tool for protein-responsive gene regulation.

References

1. Wieland, M. & Fussenegger, M. Ligand-dependent regulatory RNA parts for Synthetic Biology in eukaryotes. *Curr. Opin. Biotechnol.* **21**, 760–765 (2010).
2. Church, G. M. et al. Realizing the potential of synthetic biology. *Nat. Rev. Mol. Cell Biol.* **15**, 289–294 (2014).
3. Wang, Y.-H. et al. Synthetic biology: advancing the design of diverse genetic systems. *Annu. Rev. Chem. Biomol. Eng.* **4**, 69–102 (2013).
4. Chen, Y. Y. et al. Synthetic biology: advancing biological frontiers by building synthetic systems. *Genome Biol.* **13**, 240 (2012).
5. Mali, P. et al. Cas9 as a versatile tool for engineering biology. *Nat. Methods* **10**, 957–963 (2013).
6. McIntosh, J. A. et al. Expanding P450 catalytic reaction space through evolution and engineering. *Curr. Opin. Chem. Biol.* **19**, 126–134 (2014).
7. Wieland, M. & Fussenegger, M. Engineering molecular circuits using synthetic biology in mammalian cells. *Annu. Rev. Chem. Biomol. Eng.* **3**, 209–234 (2012).
8. Bogdanove, A. J. & Voytas, D. F. TAL effectors: customizable proteins for DNA targeting. *Science*. **333**, 1843–1846 (2011).
9. Wieland, M. & Hartig, J. S. Improved aptazyme design and in vivo screening enable riboswitching in bacteria. *Angew. Chemie* **47**, 2604–2607 (2008).
10. Galloway, K. E. et al. Dynamically reshaping signaling networks to program cell fate via genetic controllers. *Science*. **341**, 1235005 (2013).
11. Weng, D., Masci, P. & Radka, S. A phase I clinical trial of a ribozyme-based angiogenesis inhibitor targeting vascular endothelial growth factor receptor-1 for patients with refractory solid tumors. *Mol. Cancer Ther.* **4**, 948–955 (2005).
12. Win, M. N. & Smolke, C. D. A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14283–14288 (2007).
13. Vinkenborg, J. L. et al. Aptamers for allosteric regulation. *Nat. Chem. Biol.* **7**, 519–527 (2011).
14. Chang, A. L. et al. Synthetic RNA switches as a tool for temporal and spatial control over gene expression. *Curr. Opin. Biotechnol.* **23**, 679–688 (2012).

15. Uhlenbeck, O. C. A small catalytic oligoribonucleotide. *Nature* **328**, 596–600 (1987).
16. Birikh, K. R. et al. The structure, function and application of the hammerhead ribozyme. *Eur. J. Biochem.* **245**, 1–16 (1997).
17. Mathews, D. H. Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *RNA* **10**, 1178–1190 (2004).
18. Mathews, D. H. & Turner, D. H. Prediction of RNA secondary structure by free energy minimization. *Curr. Opin. Struct. Biol.* **16**, 270–278 (2006).
19. Zadeh, J. et al. NUPACK: analysis and design of nucleic acid systems. *J. Comput. Chem.* **32**, 170–173 (2011).
20. Liang, J. C. et al. A high-throughput, quantitative cell-based screen for efficient tailoring of RNA device activity. *Nucleic Acids Res.* **40**, e154 (2012).
21. Aquino-Jarquin, G. & Toscano-Garibay, J. D. RNA aptamer evolution: two decades of SELEction. *Int. J. Mol. Sci.* **12**, 9155–9171 (2011).
22. Win, M. N. & Smolke, C. D. Higher-order cellular information processing with synthetic RNA devices. *Science.* **322**, 456–460 (2008).
23. Wei, K. Y. et al. A yeast-based rapid prototype platform for gene control elements in mammalian cells. *Biotechnol. Bioeng.* **110**, 1201–1210 (2013).
24. Beisel, C. L. et al. Model-guided design of ligand-regulated RNAi for programmable control of gene expression. *Mol. Syst. Biol.* **4**, (2008).
25. Beisel, C. L. et al. Design of small molecule-responsive microRNAs based on structural requirements for Drosha processing. *Nucleic Acids Res.* **39**, 2981–2994 (2011).
26. Saito, H. et al. Synthetic human cell fate regulation by protein-driven RNA switches. *Nat. Commun.* **2**, 160 (2011).
27. Qi, L. S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
28. Daugherty, P. S. et al. Flow cytometric screening of cell-based libraries. *J. Immunol. Methods* **243**, 211–227 (2000).

29. Kosuri, S. et al. Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 14024–14029 (2013).
30. Kau, T. R. et al. Nuclear transport and cancer: from mechanism to intervention. *Nat. Rev. Cancer* **4**, 106–117 (2004).