

**Engineering RNA Devices for Gene Regulation, Biosensing,  
and Higher-Order Cellular Information Processing**

Thesis by

Maung Nyan Win

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2008

(Defended May 12, 2008)

© 2008

Maung Nyan Win

All Rights Reserved

## **Acknowledgements**

I would first and foremost like to thank my research advisor, Christina Smolke, for her continuous support and creative and insightful advice provided to me throughout my graduate study time at Caltech. Her research enthusiasm, scholarship, and dedication has inspired and challenged me as a researcher. Her guidance has been invaluable in helping me accomplish my thesis projects. It has been a great honor and a distinct pleasure to work with her at both professional and social levels.

I would also like to thank my thesis committee members, Judith Campbell, Carl Parker, and Anand Asthagiri, for their helpful comments and inputs. I have been privileged and enjoyed working in an environment that is pleasant, wonderful, and full of friendly professionals. Of course, the environment spans from our own laboratory to the Division of Chemistry and Chemical Engineering and the entire Caltech campus. I am grateful to have both professional and social support from my colleagues in the Smolke group. Thanks to Kristy Hawkins and Andrew Babiskin for frequently providing their ‘ready-made’ plasmid vectors, which saved me from slaving over a few clonings. I have also enjoyed the company of a few ‘lunch and coffee’ friends, especially, Andrew Babiskin, Joe Liang, and Leo d’Espaux; without their wonderful and fun company, the five-year period would have felt twice as long.

I would like to show my great appreciation to my cousin and her husband, Janice and David Lye. I came to the United States about nine years ago from Burma and started my undergraduate study at Virginia Commonwealth University while staying at my cousin and her husband’s house. Without their help and support, I would not have become a graduate student at Caltech. I would also like to attribute this to my undergraduate university and distinct professors, Rachel Chen, Gary Huvard, and Gary Wnek, who have laid a foundation

for my graduate study at Caltech. I would like to acknowledge my partner, Steve Fitzgerald, who has been a wonderful companion for my graduate career at Caltech, and I am very much thankful to have him as my partner.

Finally, none of my achievements would have been remotely possible without the help and support of my parents. I am deeply, deeply grateful for everything that they have done for me.

**Engineering RNA Devices for Gene Regulation, Biosensing, and  
Higher-order Cellular Information Processing**

Maung Nyan Win

B.S., Virginia Commonwealth University

M.S., California Institute of Technology

Ph.D., California Institute of Technology

**Abstract**

The proper regulation of gene expression is critical to many biological processes occurring in the cell. It is becoming increasingly apparent that post-transcriptional processing pathways play significant roles in regulating the expression of various genes in both prokaryotic and eukaryotic organisms, where they direct a variety of complex cellular functions. A striking example of a biological communication and control system directing sophisticated gene expression regulation through precise molecular recognition is the class of RNA regulatory elements, called riboswitches, comprised of distinct sensor (ligand-binding) and actuator (gene-regulatory) functions that control gene expression in response to changing levels of specific target ligand concentrations.

Inspired by these natural examples, numerous synthetic riboswitch systems have been developed and have made profound contribution to the field of riboswitch engineering. However, these early examples of synthetic riboswitches pose one or more challenges, such as portability of the switch design across different cellular systems and modularity and programmability of the components comprising the switch molecule. Therefore, we set out to develop a modular and extensible RNA-based gene-regulatory platform that will provide a

framework for the reliable design and construction of gene regulatory systems that can control the expression of specific target genes in response to effector molecules of interest. The platform is called the “ribozyme switch” and composed of distinct functional components, which are modularly coupled and functionally independent of each other. Through this platform, ribozyme switch devices that enable up- or down-regulation of target gene expression were developed. Design modularity and response programmability of the switch platform were also demonstrated. We also exhibited the versatility of the platform in implementing application-specific control systems for small molecule-mediated regulation of cell growth and non-invasive *in vivo* sensing of metabolite production.

Through the ribozyme switch platform, we further constructed higher-order RNA devices that enable complex cellular information processing operations, including logic control (AND, NOR, and NAND gates), advanced computation (bandpass filter and signal shift in the output swing), and cooperativity (signal gain). Finally, we extended the small ribozyme switch platform responsive to small molecules to a different class of ligand molecules, proteins, by developing protein-responsive gene regulators and cellular biosensors. In addition to engineering RNA devices for programming cellular function, we also developed a high-throughput method for functional characterization of small molecule-binding RNA aptamers, which enables robust, accurate, and rapid characterization of such RNA aptamers. This method can be very useful as we (and others) develop RNA aptamers for small molecules of specific interest, which can be subsequently integrated into the ribozyme switch platform as sensing elements for specific applications. Together, these research developments hold synergistic values for the reliable construction of ‘designer’ gene-regulatory systems for various biotechnological and medical applications.

## Table of Contents

<b>Acknowledgments</b>	iii
<b>Abstract</b>	v
<b>Table of Contents</b>	vii
<b>List of Tables</b>	xii
<b>List of Figures</b>	xiii
<b>Chapter I: Introduction</b>	1
<b>1.1. RNA as a versatile and powerful gene-regulatory element</b>	1
<b>1.2. RNA as sensory elements that exhibit universal sensing/binding properties</b>	3
<i>1.2.1. RNA elements that serve as thermosensors</i>	4
<i>1.2.2. RNA elements that bind nucleic acids</i>	4
<i>1.2.3. RNA elements that bind molecular ligands</i>	4
<b>1.3. Riboswitches are ligand-responsive RNA regulators of gene expression</b>	7
<i>1.3.1. General composition and conformational dynamics of riboswitches</i>	9
<i>1.3.2. Mechanisms of ligand-controlled gene regulation by riboswitches</i>	10
<i>1.3.3. Riboswitch targets and implementation in metabolic networks</i>	12
<b>1.4. Current synthetic riboswitch systems for ligand-mediated regulation of target gene expression</b>	12
<i>1.4.1. Riboswitch construction based on aptamer insertion within a transcript</i>	13
<i>1.4.2. Riboswitch construction based on direct attachment of the aptamer to a regulatory element</i>	16
<i>1.4.3. Riboswitch construction based on an evolved linker between the aptamer and regulatory domains</i>	18
<b>1.5. Further advancing the current field of engineering synthetic riboswitch systems</b>	22
<b>1.6. Interrelationship among the thesis projects</b>	23
<b>References</b>	24

<b>Chapter II: Codeine-binding RNA aptamers and rapid determination of their binding constants using a direct coupling surface plasmon resonance assay</b>	33
<b>Abstract</b>	33
<b>2.1. Introduction</b>	34
<b>2.2. Results</b>	36
2.2.1. <i>Selection of codeine-binding RNA aptamers</i>	36
2.2.2. <i>Qualitative assessment of codeine-binding affinity of the enriched final pool</i>	38
2.2.3. <i>Determination of small molecule-aptamer binding constants using a direct coupling surface plasmon resonance assay</i>	39
2.2.4. <i>Assays reveal distinct specificities of the codeine-binding aptamers to other benzyloquinoline alkaloid targets</i>	43
2.2.5. <i>Characterization of mini-aptamers that demonstrate binding affinities similar to the full-length aptamers</i>	45
2.2.6. <i>Characterization of modified mini-aptamer sequences supports the proposed secondary structures</i>	48
2.2.7. <i>Validation of the direct coupling SPR assay for characterization of small molecule-aptamer binding properties</i>	52
<b>2.3. Discussion</b>	56
<b>2.4. Materials and Methods</b>	59
2.4.1. <i>DNA template library preparation</i>	59
2.4.2. <i>Codeine coupling and affinity chromatography matrix preparation</i>	60
2.4.3. <i>Initial RNA library pool preparation</i>	60
2.4.4. <i>In vitro selection of codeine-binding aptamers</i>	61
2.4.5. <i>Aptamer library sequence analysis</i>	62
2.4.6. <i>Qualitative binding affinity assay</i>	62
2.4.7. <i>Quantitative direct coupling small molecule-aptamer binding assay</i>	63
2.4.8. <i>Isocratic affinity elution and specificity assays</i>	64
2.4.9. <i>Truncation experiments</i>	65
2.4.10. <i>Structural probing assay</i>	66



2.4.11. Dopamine aptamer binding assay	66
<b>2.5. Supplementary Information</b>	68
Supplementary Figures	68
<b>Acknowledgements</b>	70
<b>References</b>	71
<b>Chapter III: A modular and extensible RNA-based gene-regulatory platform for engineering cellular function</b>	76
<b>Abstract</b>	76
<b>3.1. Introduction</b>	77
<b>3.2. Results</b>	78
3.2.1. Component specification for a scalable and portable gene-regulatory system	78
3.2.2. Design strategies for engineering portability, utility, and composability into a biological control system	79
3.2.3. Engineering mechanisms for information transmission between the modular switch domains	82
3.2.4. Rational tuning strategies enable programming of switch regulatory response	87
3.2.5. The ribozyme switch platform exhibits component modularity and specificity	88
3.2.6. Component modularity enables implementation of ribozyme switches as regulatory systems in diverse applications	90
<b>3.3. Discussion</b>	93
<b>3.4. Materials and Methods</b>	95
3.4.1. Plasmid, switch construction, and cell strains	95
3.4.2. RNA secondary structure prediction and free energy calculation	96
3.4.3. Ribozyme characterization assays	96
3.4.4. Cell growth regulation assays	97
3.4.5. Metabolite sensing assays	97
3.4.6. Fluorescence quantification	98

3.4.7. <i>Quantification of cellular transcript levels</i>	99
<b>3.5. Supplementary Information</b>	99
Supplementary Text 3.1: <i>Glossary of terms</i>	99
Supplementary Text 3.2: <i>Ribozyme control constructs for loop sequence         coupling and stem integration controls</i>	106
Supplementary Text 3.3: <i>Rational tuning strategies for strand displacement-         based switches</i>	107
Supplementary Figures and Tables	112
<b>Acknowledgements</b>	121
<b>References</b>	122
<b>Chapter IV: Higher-order cellular information processing through synthetic RNA devices</b>	126
<b>Abstract</b>	126
<b>4.1. Introduction</b>	127
<b>4.2. Results</b>	128
4.2.1. <i>Functional RNA device composition framework and general signal         integration schemes</i>	128
4.2.2. <i>Higher-order RNA device based on SI 1         (signal integration within the 3' UTR)</i>	131
4.2.3. <i>Higher-order RNA devices based on SI 2         (signal integration at the ribozyme core)</i>	134
4.2.4. <i>Higher-order RNA devices based on SI 3         (signal integration through a single ribozyme stem)</i>	136
<b>4.3. Discussion</b>	140
<b>4.4. Materials and Methods</b>	142
4.4.1. <i>Plasmid construction, cloning, and cell strains</i>	142
4.4.2. <i>RNA secondary structure prediction, free energy calculation, and         corresponding proposed mechanism</i>	143
4.4.3. <i>In vivo assays for characterization of RNA device properties and         fluorescence quantification</i>	144

4.4.4. <i>Characterization of device higher-order information processing properties</i>	145
<b>4.5. Supplementary Information</b>	146
Supplementary Text 4.1: <i>RNA device response properties and standards in data presentation</i>	146
Supplementary Text 4.2: <i>Predicted and observed response properties of coupled single-input gates</i>	149
Supplementary Text 4.3: <i>Layered architectures extend the information processing capabilities of SI 1</i>	153
Supplementary Text 4.4: <i>Non-layered architectures (SI 2, SI 3) for an OR gate operation</i>	154
Supplementary Text 4.5: <i>Programming signal gain through multiple sensor-transmitter components</i>	156
Supplementary Text 4.6: <i>Device sequences</i>	157
Supplementary Figures and Tables	171
<b>Acknowledgements</b>	183
<b>References</b>	183
<b>Chapter V: Engineering protein-responsive gene regulators and cellular biosensors</b>	189
<b>Abstract</b>	189
<b>5.1. Introduction</b>	190
<b>5.2. Results</b>	193
5.2.1. <i>General composition framework and construction scheme for protein-responsive ribozyme switches</i>	193
5.2.2. <i>Development of protein-responsive ribozyme switches</i>	195
5.2.3. <i>In vivo functional activity of p50-responsive ribozyme switches</i>	197
<b>5.3. Discussion and Future Work</b>	199
<b>5.4. Materials and Methods</b>	200
5.4.1. <i>Plasmid construction and transformation</i>	200

5.4.2. <i>RNA secondary structure prediction and free energy calculation</i>	201
5.4.3. <i>Ribozyme characterization assays</i>	202
5.4.4. <i>Fluorescence quantification</i>	202
<b>Acknowledgements</b>	203
<b>References</b>	203
<b>Conclusions</b>	206

## List of Tables

- Table 2.1. Codeine-binding affinities of the full-length aptamer sequences as determined from the direct coupling SPR assay.
- Table 2.2. Dissociation rate constants ( $k_{\text{off}}$ ) for codeine binding of the final pool, FC5, FC45, and their corresponding truncated sequences. The corresponding association rate constant ( $k_{\text{on}}$ ) is equivalent to  $k_{\text{off}}/K_d$ .
- Table 2.3. Codeine-binding affinities of several full-length aptamers determined from replicate SPR binding assays for method reproducibility assessment.
- 
- Supplementary Table 3.1. Relative steady-state ribozyme switch and ribozyme control transcript levels in the presence and absence of theophylline.
- Supplementary Table 3.2. Free energies ( $-\Delta G$ , kcal/mol) of individual conformations (ribozyme-active and -inactive) and the energy difference ( $\Delta\Delta G$ , kcal/mol) between the free energies of these two conformations predicted by RNAstructure 4.2.
- Supplementary Table 3.3. Primer and additional ribozyme construct sequences.
- Supplementary Table 4.1. The basal output signals and output swings of the RNA devices studied in this work are shown in % device response over the full transcriptional range of the employed promoter.
- Supplementary Table 4.2. Free energy changes associated with RNA devices comprised of internal Buffer and Inverter gates and associated Hill coefficients. Free energy changes between RNA device states are predicted from a standard RNA folding program, RNAstructure 4.2.

## List of Figures and Supplementary Figures

- Figure 1.1. A schematic illustration of an *in vitro* selection process known as SELEX.
- Figure 1.2. A schematic diagram of a typical riboswitch composed of two distinct domains.
- Figure 1.3. A schematic illustration of mechanisms by which riboswitches achieve gene expression regulation in response to their target metabolite binding.
- Figure 1.4. A schematic illustration of riboswitches constructed based on aptamer insertion within a target transcript.
- Figure 1.5. A schematic illustration of riboswitches constructed based on direct attachment between the aptamer and regulatory domains.
- Figure 1.6. A schematic illustration of riboswitches constructed based on an evolved linker between the aptamer and regulatory domains.
- Figure 2.1. Codeine-binding RNA aptamer clone sequences.
- Figure 2.2. Schematics of the codeine-immobilized surfaces used in the *in vitro* selection process and SPR binding property assay.
- Figure 2.3. Concentration-dependent codeine-binding responses and the corresponding equilibrium binding curve.
- Figure 2.4. Equilibrium codeine-binding response curves.
- Figure 2.5. The FC5 and FC45 aptamers exhibit differing specificities to BIA structural analogues.
- Figure 2.6. Codeine-binding mini-aptamer characterization.
- Figure 2.7. Structural stabilization and sequence requirements of the FC5 mini-aptamer stems.
- Figure 2.8. Structural stabilization of the FC45 mini-aptamer.
- Figure 2.9. Validation of the direct coupling SPR assay.
- Figure 3.1. General design strategy for engineering ribozyme switches.
- Figure 3.2. Regulatory properties of the strand displacement information transmission mechanism.
- Figure 3.3. Regulatory properties of the helix slipping information transmission mechanism
- Figure 3.4. Tunability of the strand displacement-based ribozyme switches.
- Figure 3.5. Modularity and specificity of the strand displacement-based ribozyme switches.

- Figure 3.6. System modularity of ribozyme switches enables implementation in programmed cell growth.
- Figure 3.7. System modularity of ribozyme switches enables implementation in non-invasive detection of metabolite biosynthesis.
- Figure 4.1. Functional RNA device composition framework.
- Figure 4.2. Higher-order RNA devices based on signal integration within the 3' UTR (SI 1).
- Figure 4.3. Higher-order RNA devices based on signal integration at the ribozyme core (SI 2).
- Figure 4.4. Higher-order RNA devices based on signal integration at a single ribozyme stem (SI 3).
- Figure 5.1. General composition framework and modular design strategy for engineering ligand-controlled ribozyme switch-based gene regulatory systems
- Figure 5.2. A general construction scheme for protein-responsive ribozyme switches. Color schemes follow those described in Figure 5.1.
- Figure 5.3. Sequences and structures of RNA aptamers with affinity and specificity for NF- $\kappa$ B p50.
- Figure 5.4. Modular design strategies and systematic engineering of an NF- $\kappa$ B p50-responsive ribozyme switch.
- Figure 5.5. Sequences and secondary structures of two p50-responsive ribozyme switches in their ligand-bound conformations.
- Figure 5.6. The device response of p50-responsive ribozyme switches exhibiting ON switch regulatory responses.

Supplementary Text 4.2 Figure 1. The device response over varying input concentrations of representative coupled gate devices (2xL2bulgeOff1, right; 2xL2bulge1, left) constructed through SI 1.1 and their corresponding single-gate device counterparts (L2bulgeOff1, L2bulge1).

Supplementary Text 4.3 Figure 1. Schematic representation of layered architectures that extend the information processing capabilities of SI 1. Left, schematic illustrating a NAND gate operation by

inverting the output of an AND gate. Right, schematic illustrating an OR gate operation by inverting the output of a NOR gate.

Supplementary Text 4.4 Figure 1. Schematic representation of an RNA device based on SI 2 that functions as an OR gate operator.

Supplementary Text 4.4 Figure 2. OR gate devices.

Supplementary Figure 2.1. Qualitative assessment of the enrichment in codeine-binding affinity of the final aptamer pool.

Supplementary Figure 2.2. Alternative FC45 mini-aptamer structures support the proposed structure of the FC45 mini-aptamer.

Supplementary Figure 2.3. Structural stabilization studies of FC5 and FC45 mini-aptamers.

Supplementary Figure 2.4. Structural probing results through lead-induced and RNase T1 cleavage patterns.

Supplementary Figure 3.1. Control constructs supporting the design strategy for engineering ligand-regulated ribozyme switches.

Supplementary Figure 3.2. Flow cytometry histograms of L2bulge1, L2bulgeOff1, and the ribozyme control cell populations grown in the presence (+) and absence (-) of 5 mM theophylline.

Supplementary Figure 3.3. Flow cytometry histograms of the helix slipping-based ribozyme switch cell populations grown in the presence (+) and absence (-) of 5 mM theophylline.

Supplementary Figure 3.4. Temporal responses of L2bulge1, L1cm10, and L2cm4 in response to the addition of 5 mM theophylline (final concentration).

Supplementary Figure 3.5. Sequences and structures of tuned ribozyme switches in the L2bulge series.

Supplementary Figure 3.6. Flow cytometry histograms of the tuned ribozyme switch series cell populations grown in the presence (+) and absence (-) of 5 mM theophylline.



- Supplementary Figure 3.7. Detection of intracellular accumulation of the substrate xanthosine and the product xanthine over three different time points.
- Supplementary Figure 3.8. Dynamic ranges of regulation of the ribozyme switches and controls engineered in this work.
- Supplementary Figure 4.1. The device response and truth table of an AND gate operator (L2bulge9+L2bulge1tc) based on SI 1.2.
- Supplementary Figure 4.3. The device response and truth table of a NOR gate operator (L2bulgeOff1+L2bulgeOff2tc) based on SI 1.3.
- Supplementary Figure 4.4. Schematic representation and device response of a bandpass filter operator (L2bulge1+L2bulgeOff1) based on SI 1.4.
- Supplementary Figure 4.5. The device response and truth table of a NAND gate operator (L1cm10-L2bulgeOff1tc) based on SI 2.1.
- Supplementary Figure 4.6. The device response and truth table of AND gate operators (tc-theo-On2 and tc-theo-On3) based on SI 3.1.
- Supplementary Figure 4.7. The device response of RNA devices comprised of internal Buffer and Inverter gates and their single internal gate device counterpart (L2bulge1).
- Supplementary Figure 4.8. The device response over varying input concentrations of a representative RNA device comprised of internal Buffer and Inverter gates (theo-theo-On1) and its single internal gate device counterpart (L2bulge1) demonstrates no signal gain ( $n_H \approx 1$ ).
- Supplementary Figure 4.9. The device response over varying input concentrations of RNA devices comprised of internal Buffer and Inverter gates (theo-theo-On10–12) and their single internal gate device counterpart (L2bulge1) demonstrates programmed cooperativity.
- Supplementary Figure 4.10. The device response of RNA devices comprised of two internal Inverter gates and their single internal gate device counterpart (L2bulgeOff1).

- Supplementary Figure 4.11. The device response over varying theophylline concentrations of representative RNA devices comprised of two internal Inverter gates (theo-theo-Off2,  $n_H \approx 1$ ; theo-theo-Off6,  $n_H \approx 1.2$ ), and their single internal gate device counterpart (L2bulgeOff1,  $n_H \approx 1$ ).
- Supplementary Figure 4.12. The device response of a representative RNA device comprised of internal Buffer and Inverter gates (theo-theo-On1) and its mutated sensor variants demonstrates that input binding at both internal gates is responsible for the device response.
- Supplementary Figure 4.13. The device response of a representative RNA device comprised of two internal Inverter gates (theo-theo-Off2) and its mutated sensor variants demonstrates that input binding at both internal gates is responsible for the device response.
- Supplementary Figure 4.14. The device response of a representative RNA device comprised of internal Buffer and Inverter gates that exhibits programmed cooperativity (theo-theo-On13) and its mutated sensor variants demonstrates that input binding at both internal gates is responsible for the device response.
- Supplementary Figure 4.15. The device response of a representative RNA device comprised of two internal Inverter gates that exhibits programmed cooperativity (theo-theo-Off6) and its mutated sensor variants demonstrates that input binding at both internal gates is responsible for the device response.
- Supplementary Figure 4.16. Secondary structures and sequences of input-bound states of representative RNA devices.
- Supplementary Figure 4.17. Secondary structures and sequences of input-bound states of representative RNA device comprised of internal Buffer and Inverter gates responsive to the same input, illustrating points of coupling of two sensor-transmitter components.