RNA-Based Computing Devices for Intracellular and Diagnostic Applications

by

Duo Ma

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2019 by the
Graduate Supervisory Committee:

Alexander Green, Chair
Yan Liu
Marco Mangone

ARIZONA STATE UNIVERSITY

August 2019

ABSTRACT

The fundamental building blocks for constructing complex synthetic gene networks are effective biological parts with wide dynamic range, low crosstalk, and modularity. RNA-based components are promising sources of such parts since they can provide regulation at the level of transcription and translation and their predictable base pairing properties enable large libraries to be generated through in silico design. This dissertation studies two different approaches for initiating interactions between RNA molecules to implement RNA-based components that achieve translational regulation. First, single-stranded domains known as toeholds were employed for detection of the highly prevalent foodborne pathogen norovirus. Toehold switch riboregulators activated by trigger RNAs from the norovirus RNA genome are designed, validated, and coupled with paper-based cell-free transcription-translation systems. Integration of paper-based reactions with synbody enrichment and isothermal RNA amplification enables as few as 160 copies/mL of norovirus from clinical samples to be detected in reactions that do not require sophisticated equipment and can be read directly by eye. Second, a new type of riboregulator that initiates RNA-RNA interactions through the loop portions of RNA stem-loop structures was developed. These loop-initiated RNA activators (LIRAs) provide multiple advantages compared to toehold-based riboregulators, exhibiting ultralow signal leakage in vivo, lacking any trigger RNA sequence constraints, and appending no additional residues to the output protein. Harnessing LIRAs as modular parts, logic gates that exploit loop-mediated control of mRNA folding state to implement AND and OR operations with up to three sequence-independent input RNAs were constructed. LIRA circuits can also be ported to paper-based cell-free reactions to implement portable systems with molecular computing and sensing capabilities. LIRAs can detect RNAs from a variety of different pathogens, such as HIV, Zika, dengue, yellow fever, and norovirus, and after coupling to isothermal amplification reactions, provide visible test results down to concentrations of 20 aM (12 RNA copies/μL). And the logic functionality of LIRA circuits can be used to specifically identify different HIV strains and influenza A subtypes. These findings demonstrate that toehold- and loop-mediated RNA-RNA interactions are both powerful strategies for implementing RNA-based computing systems for intracellular and diagnostic applications.

DEDICATION

To my father, Guangtian, who helps me find the direction in my life.

To my mother, Chuanling, who encourages me and makes me feel warm inside.

To my wife, Yueming, who gives me happiness and supports me all the time.

To my son, Max, who can run by himself now and interested in everything around.

To my family, who support me all the time.

ACKNOWLEDGMENTS

I would like to express my deepest respect and appreciation to my advisor, Dr. Alexander A. Green, for his extraordinary guidance and support during my graduate research and study at Arizona State University. Without his mentorship, it would be impossible for me to accomplish all that I have. And I will try my best in my future life to be a researcher like him, who can always think out creative ideas and keep an energetic attitude toward science.

I would also like to thank my committee members, Dr. Yan Liu and Dr. Marco Mangone, for their helpful guidance on my research and future developments over the years, and for the knowledge I learned from their classes, which makes me well prepared for my graduate researches.

I thank the Center for Molecular design and biomimicry and the Biodesign Institute at Arizona State University for providing the opportunity to work with an excellent group of scientists and advanced facilities.

I would like to thank all my current and former lab mates and friends, Anli Tang, Abhishek Debnath, Ahmed Yousaf, Yuexin Li, Soma Chaudhary, Kaiyue Wu, Zhaoqing Yan, Matthew Gilliam, Sanchari Saha and Griffin McCutcheon, for keeping me accompanied and helping me through the PhD life.

I am grateful to the FD3000 Sponsored Funds, GR16137 LOW-COST NUCLEIC ACID DIAGNOSTIC, for supporting of my research.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Synthetic biology

Synthetic biology is a highly interdisciplinary research field that integrates our current understanding of several subjects, such as biology, chemistry and engineering. Biology studies provide us with deeper understanding of natural systems, from which we can learn the basic mechanisms that hide behind biological interactions and develop new ideas to engineer natural systems by integrating them with engineering approaches. Novel molecules and molecular systems that are required for carrying out modifications in synthetic biology can be synthesized by chemistry[1]. The term 'synthetic biology' was first used by Barbara Hobom in 1980 for bacteria that had been engineered via recombinant DNA technology[2,3]. Indeed, the development of genetic engineering technologies paved the way for the development of synthetic biology. The advent of genetic engineering dates back to the 1960s, when the study of *lac* operon in E. coli demonstrated the regulatory circuits exist in natural systems[4]. Later in the 1970s and 1980s, genetic manipulation techniques, such as molecular cloning and PCR techniques, gradually became more and more widely used in molecular biology research[5]. During the 1990s, automated DNA sequencing and high-throughput analysis of RNA, protein and other cellular components were developed. These large data-based studies drove researchers to analyze organisms in an entirely way and study biological networks in a 'top-down' way. Those studies were then classified as a new discipline, termed 'systems biology'[6-8]. From systems biology, researchers gradually made it clear that how different biological parts function in natural systems, and what influence would they bring to natural systems by making modifications to those biological parts. With the development of our understanding of versatile biological parts, a 'bottom-up' approach for biological network construction was gradually developed, in which well-defined biological parts attained from systems biology studies were constructed into complex biological circuits to fulfill artificial functions. This discipline was then termed as 'synthetic biology'.

The goal of synthetic biology is to endow organisms with the ability to carry out new biological functions. Synthetic biologists tend to take computer hardware engineering as an

example, in which transistors, capacitors and resistors work as the bottom layer, which could be assembled into gates, modules and finally into computers. Computers in different locations can contact each other by networks. Similarly, in synthetic biology, starting from basic biological materials, such as DNA, RNA and protein, devices with a single or multiple predefined functions are first designed. Multiple devices that could implement different functions work together to generate versatile biological circuits, which then bring new biological functions to host cells. To carry out more sophisticated tasks that cannot be done by a single cell population, multiple cell populations can be engineered to communicate and coordinate with each other. In such systems, different cell populations need to be endowed with different functions that are required in the targeting coordinate tasks. Those cell populations can be 'wired' together with predefined biological circuits, using small molecules or proteins as signals for transmission[8]. Thus, no matter in single-cell or cell-population engineering, biological circuits with reliable synthetic function represents fundamental building blocks in synthetic biology.

For constructing artificial biological circuits, synthetic biologists adopt two different approaches. One is by using unnatural molecules to reproduce biological behavior, and the other one is by assembling biological parts into artificial biological circuits[2]. However, unlike electric circuits, in which different devices and modules are well insulated, all biological components are randomly distributed in cells and affected by cellular environment. Considering our incomplete understanding of biological systems, the assembly of parts into complex synthetic circuits is still hindered by unknown interactions in cellular environment. Moreover, the complexity of part coordination will increase exponentially with the increasing number of parts being incorporated into circuits[9]. Last but not least, engineered biological devices and circuits that function well in one type of cellular context may not work well or may even be toxic when transferred into other types of cells or organisms[9].

Although a lot of obstacles obstruct the ability to engineer versatile biological devices and circuits, researchers have developed numerous biological devices and circuits with complex functions in the past two decades. In DNA level designs, genetic circuits were successfully constructed, including genetic toggle switch[10], synthetic oscillator[11-14], recombinase genetic

switch[15] and derived logic gates[16-18], counters[19], promoter regulation based logic circuits[20-24], orthogonal ribosome based logic circuits[25], noise propagation cascade[26], band-pass filter[27], E. coli based optical sensors[28,29] and time-delay circuits[30], etc. For RNA-level constructs, lots of designs were also successfully tested, like ligand-responsive riboregulator[31] and derived logic gates[32], RNAi based genetic switch[33] and derived logic circuits[34], riboswitches[35-45], ligand-responsive ribozyme[46], engineered riboregulator[47], engineered ribosome-mRNA pairs[48], small transcription activating RNAs[49] and transcriptional activators[50,51], etc. With respect to protein-level, biological circuits with high efficiency were also constructed, although not as many as DNA or RNA based designs, such as protein based allosteric gating[52], protein receptors for non-native target ligands[53], protein-binding domain based logic gates[54], transcription effector based logic gates[55] and proteolysis-based logic circuits[56], etc. Even for synthetic pathway construction and multicellular communication, researchers also obtained great breakthroughs, such as the constructions of butanol producing pathway[57], artemisinic acid producing pathway[58] and fuels and chemicals producing pathways[59-61], as well as multicellular systems, like pulse generator[62], programmed pattern formation[63], programmed population control[64], analog-to-digital converter[65], programmable full-adder[66], etc.

The above circuit classifications are not quite stringent, since every circuit requires incorporation of all cellular components, including DNA, RNA, proteins, etc., from host cells to carry out their functions. The rough classification only considered the engineered part that plays a more important role in their designed biological functions. But from the biological circuits listed above, we can clearly see that the studies of DNA and RNA based circuits have advanced more than protein, pathway and multicellular studies. This is likely the result of the ease of engineering DNA and RNA compared with relatively higher difficulty for protein, pathway and multicellular engineering, considering our limited understanding of biological mechanisms. DNA or RNA based researches have more clues for people to follow. Versatile molecular biological techniques have been developed for researchers to modify or to engineer DNA and RNA outside or inside the cells. Unlike protein, which consists of 20 amino acids and unpredictable structures, DNA and RNA only contains four bases with predictable Watson-Crick base pairing rules, whose structure is

more predictable. Using DNA and RNA sequencing techniques, DNA or RNA sequences that have biological functions can be well identified, which facilitate the modification and reconstitution of their biological functions. With the predictable base pairing rules and well defined biological functions, engineered DNA or RNA molecules with prescribed functions and predefined structures can be well designed, which lessens the burden of devices screening and tuning. For the enumerated DNA studies, most of them rely on protein controlled gene transcription in cells, either with transcription factors or repressors.

1.2 General genetic circuits

In the year of 2000, the development of toggle switch design led to a breakthrough in synthetic biology studies. The toggle switch is a DNA sequence containing multiple domains with biological functions, which consists of two repressor genes regulated by two constitutive promoters[10]. Each repressor gene encodes for a repressor protein, which can inhibit the transcription of the other repressor gene by inhibiting the binding of RNA polymerase with its corresponding constitutive promoter (Fig. 1.1A). The mutual inhibition of those two repressors leads to bi-stable behavior that is without inducer, either repressor 1 or repressor 2 induced by promoter 1 or promoter 2 can be expressed predominantly in the cells transformed with an engineered plasmid containing the toggle switch DNA construct. The current active promoter can be switched to the opposite promoter by adding the inducer to deactivate current active repressor proteins, which permits the expression of opposite repressor proteins to inhibit current active promoter.

In the actual design, two pairs of promoter and repressor were selected. One pair is Ptrc-2 promoter (promoter 2) and Lac repressor (lacI, repressor 2), which can be inhibited by inducer IPTG, and the other pair is P_Ls1con promoter (promoter 1) and a temperature-sensitive λ repressor (cIts, repressor 1), which can be inhibited by elevating temperature (Fig. 1.1A). Reporter protein is GFP, which is also regulated by promoter 2. In principle, with the addition of IPTG, expression of genes regulated by promoter 2, including repressor 1 and GFP genes, can be induced, so the fluorescence signal can be detected at this stage (Fig. 1.1B). The expression

of repressor 1 protein will block the binding of RNA polymerase and promoter 1, so expression of repressor 2 can also be inhibited. When IPTG was removed from the culture medium, since no repressor 2 protein is synthesized by the host cell, promoter 2 will remain active, which keeps inducing the expression of GFP and repressor 1 proteins. However, when temperature increases, repressor 1 proteins will be deactivated and cannot inhibit the function of promoter 1, thus repressor 2 protein regulated by promoter 1 can start to express to inhibit promoter 2, which terminates the expression of repressor 1 and GFP reporter proteins. In this case, a decreasing fluorescence signal should be detected.

Oscillator systems are another example of using promoter-repressor pairs[11]. In this design, three promoter-repressor pairs were constructed into one plasmid, including $P_L$lac promoter and *lacI* repressor, $P_L$tet promoter and *tetR* repressor, λP_R promoter and *λcI* repressor (Fig. 1.2A). The first repressor protein, *lacI* blocks the binding between RNA polymerase and $P_L$lac promoter, so the expression of *lacI* will inhibit the expression of the second repressor, tetR. And tetR inhibits the interaction of RNA polymerase with $P_L$tet promoter, and thus regulate the expression of the third repressor *cI*, which is arranged to inhibit the expression of *lacI* repressor protein to complete the cycle by inhibiting λP_R promoter. The reporter protein appears on a separate plasmid, whose expression is inhibited by tetR repressor. A transient pulse of IPTG will deactivate the function of *lacI* repressor protein, which can induce the expression of *tetR* repressor. The expression of *tetR* repressor will then inhibit the expression of both *cI* repressor and GFP protein. Thus, by adding IPTG, GFP signal strength will decrease initially. While the expression of *cI* repressor is repressed, λP_R promoter will be activated, so the expression of *lacI* repressor will be turned on. Once the expressed *lacI* repressor compete the transiently inputted IPTG pulse, it will inhibit the binding of RNA polymerase with $P_L$lac promoter again, and the expression of *tetR* repressor will then be repressed, and GFP expression can be turned on. This three pairs of promoter-repressor construction thus leads to an oscillatory expression of GFP protein, which is reflected in fluorescence strength output (Fig. 1.2B).

A counter was developed in a different way, which takes advantage of transcriptional cascade[19]. Two different riboregulated transcriptional counters (RTC) were developed, with one

can count to two, and the other one can count to three (Fig. 1.3). For the RTC two-counter, constitutive promoter $P_{Ltet0-1}$ drives the transcription of T7 RNA polymerase gene, which is encoding for T7 RNA polymerase. And transcription of GFP protein gene is carried out by T7 RNA polymerase (Fig. 1.3A). However, the transcribed T7 RNA polymerase and GFP protein mRNAs are regulated by riboregulators, which will inhibit the translation process by blocking RBS (ribosome binding site) to interrupt the binding of ribosome with mRNAs. To turn on the translation of these riboregulated mRNAs, a trans-activating RNA (taRNA) is designed to contain sequence that is complementary to the riboregulators, so once taRNA is transcribed by adding inducer arabinose, the structure of riboregulator will be disrupted, and the expression of T7 RNA polymerase and GFP proteins will also be turned on. Before adding arabinose, repressed T7 RNA polymerase mRNAs are automatically synthesized by host cells with the constitutive promoter. After the first pulse of arabinose, a short burst of taRNA can be transcribed, which will then disrupt the riboregulator of the repressed T7 RNA polymerase mRNA to turn on the T7 RNA polymerase expression. Then the arabinose was removed from the culture medium by centrifuge and wash, and remained intracellular arabinose and synthesized taRNA are gradually metabolized from host cells. In this case, T7 RNA polymerase that had been synthesized inside the cell can transcribe riboregulated GFP mRNA. Until the next arabinose pulse to induce the transcription of another burst of taRNA, GFP expression will remain at OFF state (Fig. 1.3B). A RTC three-counter was constructed in the same way, by inserting the T3 RNA polymerase synthesis in the middle (Fig. 1.3C&D).

Logic gate is another hot topic research field in synthetic biology, and most of the constructed logic gate circuits usually take advantage of interactions between protein, DNA and RNA. Layered designs are typically used for constructing complex circuits, in which outputs of one layer are designed into inputs for next layers. As shown in Figure 1.4, the two inputs of the whole logic circuit are two small molecules, and the outputs of the logic circuit are fluorescent proteins. Logic computation occurs in the processing unit, lies in the middle of the logic circuit. In the transcriptional control process, input molecules 1 and 2 can abolish transcription factors TF_1 and TF_2 respectively, which can interact with their corresponding promoters to activate the

transcription of downstream DNAs into mRNAs. TF1 can activate the transcription of RNA-binding proteins, and TF2 can activate the transcription of a mRNA containing a regulatory part (BOX) and reporter protein gene. In the translational control level, the expressed RNA-binding protein binds with the regulatory part of the mRNA activated by TF2 to inhibit the translation of reporter proteins, so for overall circuit, it forms a N-IMPLY logic. To turn on the fluorescence output signal, molecule 1 is required to be added into the system to abolish TF1, so that the RNA-binding protein cannot be expressed. Molecule 2 is required not to be added, so the transcription of mRNA can be successfully processed by TF2 inducement.

The regulation mechanisms for the above DNA based genetic circuits are usually designed at transcriptional level, with transcription factor and repressor proteins as major acting components. However, with the development of RNA biology, more and more natural RNA-based regulatory systems have been discovered in natural systems either in transcriptional and translational levels, which demonstrated an important role of RNA level regulations.

1.3 RNA based regulation of gene expression

From its position at the center of central dogma, RNA plays an important role in regulation of gene expression in both prokaryotic and eukaryotic cells. Antisense RNAs[67,68], small interfering RNA (siRNA)[69-71] and MicroRNA (miRNA)[72-74] can repress gene expression by binding with specifically targeting sequences in mRNAs. Riboregulator RNAs[75,76] can cause conformational change of mRNAs to turn on or turn off gene expression by binding with targeting domains. Ribozymes[77-79] located in mRNAs can affect gene expression by self-cleavage. Riboswitches, which can bind small molecules to regulate gene expression, are also discovered and have been well studied.

1.3.1 Riboswitch

A riboswitch is a non-coding RNA located in the 5' untranslated region of mRNAs, which can bind with small ligand molecules, such as nucleotides[42,45], amino acids[43] and other small molecules[39,40,44]. Binding with ligand molecule will cause conformational change of riboswitch

RNA, which then block or release RNA domains with biological functions to carry out downstream interactions. Riboswitch can either sequester or release RBS domain to regulate gene expression at translational level by binding with small molecules, or regulate gene expression at transcriptional level by disrupting the formation of transcriptional terminators.

Thiamine pyrophosphate (TPP) binding riboswitch is a widely distributed riboswitch, which can regulate gene expression in both transcriptional and translational levels[44]. TPP is an active form of vitamin B1, which plays an important role in protein-catalyzed reactions. And as a metabolite-sensing RNA regulatory system, TPP-binding riboswitches are widely distributed in all species, which regulate the importation or synthesis of thiamine and its derivatives. As shown in Figure 1.5A, with higher TPP concentration, the structure of riboswitch will stay in a condition, in which the Shine-Dalgarno (SD) sequence is sequestered in a hairpin structure. In this case, expression of downstream protein is repressed. When the concentration of TPP decreases below the threshold, a conformational change of riboswitch structure will occur, and the sequestered SD domain will also be released to a linear structure. In this condition, ribosome can recognize the SD domain to translate downstream gene into proteins.

With respect to transcriptional level control, the conformational change of riboswitch can regulate the formation of transcriptional terminators. With TPP binding, riboswitch structure can be stably formed together with the transcriptional terminator (Fig. 1.5B), which can terminate transcription to generate truncated mRNAs. But when TPP is removed, the interaction between riboswitch and transcription terminator will disrupt both of their structures, so that full-length mRNA can be successfully transcribed, and gene expression is thus tuned on.

Although riboswitches are widely distributed in nature, they are very sensitive to sequence changes, since their sequences are derived from natural evolution. Thus, regulatory systems that can be engineered are more useful for constructing more complex biological circuits to carry out more versatile biological functions.

1.3.2 RNA-only regulatory systems

Besides riboswitches which can regulate gene expression by binding with ligand molecules, RNA-only regulatory elements were also discovered in natural systems, such as antisense-RNA[80]. Antisense-RNA regulates gene expression by binding with its complementary sequences located in mRNAs. Binding of antisense-RNA will cause conformational changes in the target mRNA, which then exhibit different biological functions. In transcriptional-level regulation, antisense-RNA can regulate the formation of transcriptional terminator to regulate gene expression (Fig. 1.6A). Without antisense-RNA binding, the terminator sequence will be disrupted in a large stem-loop structure during transcription, so no transcription termination can occur. When antisense-RNA binds with its targeting sequence from mRNA, the large stem-loop structure will not form during transcription, which facilitates the formation of transcriptional terminator to terminate gene transcription. At translational-level, gene regulations are occurred at RBS domain[68,81]. Without antisense-RNA, RBS domain will stay in a linear structure, which is accessible to the binding of ribosomes (Fig. 1.6B). But when antisense-RNA binds with its target, the RBS domain will be sequestered into a hairpin structure, and gene expression is also being blocked. Another mechanism for antisense-RNA regulation is by directly binding with RBS domain through loop-linear interactions (Fig. 1.6C).

Inspired by those natural regulatory systems, researchers designed and engineered different RNA-based regulatory systems that can function in either transcriptional or translational levels.

1.3.2.1 Transcriptional level designs

By engineering the antisense-RNA and its target sequence[82], researchers got different RNA-based regulatory elements (Fig. 1.6D&E). But these mutated types usually exhibit less regulating efficiency compared with their wild-type versions, which can exhibit around 5-fold change in gene expression levels. To improve the regulating efficiency, Meyer *et al.* developed one type of small transcription activating RNAs (STARs) that directly bind with terminator sequences to disrupt its structure during transcription[49] (Fig. 1.7A). And with this design, the fold change of gene regulation can reach up to about 15-fold. Westbrook *et al.* further improved the

design into a dual-level regulation systems, in which regulation occurs at both transcriptional and translational levels[83] (Fig. 1.7B). With this updated regulating mechanism, the fold change of the system is elevated to over 900-fold.

However, although the transcriptional designs show great improvement in fold change of gene expression, a fundamental limitation of these designs is the sequence constraints of their antisense-RNAs, since those RNAs need to contain complementary sequences with transcriptional terminators, whose sequences are completely predefined. This is an avoidable obstacle for engineering those systems into more complex biological circuits.

1.3.2.2 Translational level designs

Inspired by natural translation regulatory systems, Isaacs *et al.* developed an engineered riboregulator[47] (Fig. 1.8A). In the design, RBS domain is blocked in a hairpin structure, which can prevent it from binding with ribosome, so the system is in its 'OFF' state. When the trans-activating RNA (taRNA) is induced, the hairpin structure will be disrupted, and RBS domain can also be released. In this case, the ribosome can bind with the RBS domain to translate the downstream gene into protein. This protein expressing state is termed the 'ON' state. However, several limitations of this design hindered its further applications. First limitation is the sequence constraints in the taRNA. Although this design is the first engineered riboregulator, it still cannot be engineered with completely random sequences for raRNAs, since there are two main sequence constraints located in their binding domains. One is the RBS sequence in the stem, the other one is the YUNR motif (pYrimidine-Uracil-Nucleotide-puRine) located in the hairpin loop. This in turn restricted the sequence space of the taRNA designs, which hindered its applications into constructing more complex biological circuits or in RNA detections. Second, binding between crRNA and taRNA are initiated by the binding of taRNA with the short loop from the crRNA. Since the loop size in this design is very short, only 6 nucleotides (6 nts) long, the driving force of their binding is quite weak, which resulted in low dynamic range. The best design exhibit less than 15-fold ON/OFF ratio, weaker than the transcriptional designs talked above.

To overcome those limitations, Green *et al.* designed a new kind of riboregulator, named the toehold switch[84]. In this design, RBS domain is moved to the top loop of a hairpin structure (Fig. 1.8B), and start codon is also blocked in the hairpin. Although the start codon locates in the stem part, which is covered by binding domain, a large bulge is designed in its complementary strand, so there are completely no sequence constraints for trigger RNA designs. The binding between switch and trigger RNAs are initiated by a linear-linear interaction occurred from the binding between the trigger RNA and the toehold part of switch RNA. Considering the longer length of the toehold part, which is over 15 nts long, the driving force for their binding is enormously improved. All these improvements lead to extremely high dynamic range for toehold switches, with the best one shows ON/OFF ratio around 660-fold. For different toehold switches, since there is no conserved domain in trigger RNAs, the crosstalk level between different devices is also extremely decreased compared with the conventional designs.

Toehold switch can also be engineered into sophisticated ribocomputing circuits[85]. By arrange different toehold switches in a row, a 6-input 'Or' logic nanodevice is successfully constructed (Fig. 1.8C). To carry out 'And' logic computation, the trigger RNA is split into few parts with complementary domains added onto each of them (Fig. 1.8D). When being transcribed together, those split triggers can bind together to form the full-length trigger to turn on the toehold switch. With this mechanism, a 4-input 'And' logic nanodevice is successfully tested. The most complex circuit constructed in this work is a 12-input DNF ribocomputing circuit (Fig. 1.8E). In this circuit, all 28 input conditions are successfully tested with clear signal differences between TRUE and FALSE states (Fig. 1.8F).

1.4 In vitro diagnostic

Since trigger RNAs of toehold switch do not contain any sequence constraints, toehold switches can be engineered to detect RNA molecules with completely random sequences. This leads to a major application of toehold switch, diagnostics. To get a portable and low-cost diagnostic platform based on toehold switch, researchers incorporated cell-free system[86] onto a paper-based platform[87] to carry out diagnostic tests[88].

1.4.1 Cell-free system and paper-based platform

For cell-free system, all components, including enzymes, materials and energy source, which are used for transcription and translation, are either purified individually from bacteria or directly added into the system[86]. Systems of the former type include 32 components that are purified individually: three initiation factors (IF1, IF2, IF3), three elongation factors (EF-G, EF-Tu, EF-Ts), three release factors (RF1, RF3, RRF), 20 aminoacyl-tRNA synthetases (ARSs), methionyl-tRNA transformylase (MTF), T7 RNA polymerase, and ribosomes. In addition, some other necessary components are also added into the system, including 46 tRNAs, NTPs, creatine phosphate, 10-formyl-5,6,7,8-tetrahydrofolic acid, 20 amino acids, creatine kinase, myokinase, nucleoside-diphosphate kinase, and pyrophosphatase. When DNA and RNA molecules encoding for proteins are added into the system, expression of the proteins can be transferred from inside the cells to outside the cells. Paper-based platform is a small piece of filter paper disc, which is 2 mm in diameter. Cell-free system can be freeze-dried onto the paper disk, and remain stable for years. Once rehydrated with water, transcription and translation of input DNA or RNA will be activated.

For paper-based diagnostic platform (Fig. 1.9A), the reporter gene that regulated by toehold switch is replaced with lacZ gene, which encodes for an enzyme that can cleave its chemical substrate to output a color change signal from yellow to purple. Cell-free systems together with plasmid encoding for pathogen detecting toehold switches are freeze-dried onto the small paper discs. When the freeze-dried paper disc is rehydrated with corresponding pathogen RNA, expression of the regulated lacZ protein will be activated to output a color change signal for detection. This platform exhibits several advantages. It does not require thermal cycler for running the reaction, and the color change signal can be directly discerned by eye, which does not require additional equipment for detecting the readout. Moreover, only targeting RNA sequence can turn on the color change reaction, which wipes out the possibility of getting false positive results.

1.4.2 Zika virus detection

The recent outbreak of Zika illustrates the need for low-cost and portable diagnostic platforms, which do not require expertise to manipulate and expensive instruments to implement for people who are living in rural areas to use. The traditional ways to detect Zika virus are antibody detection, qRT-PCR and isothermal nucleic acid amplification. Antibody detection usually outputs results very fast, but it provides limited sensitivity. Moreover, the cross-reactivity from other flaviviruses may result in false positive readouts[89]. qRT-PCR and isothermal nucleic acid amplification provide with more accurate diagnostic readouts since they are targeting at specific regions of virus genomes. However, these techniques require expensive equipment, reagents and experienced expertise to run the test, which may not available for people living in low-resource locations.

Paper-based diagnostic platforms provide a better pathogen detecting method. However, viruses in clinical urine or serum samples are usually at very low concentration, which is not sufficient to turn on paper-based reactions. To improve platform sensitivity, an isothermal RNA amplification technique termed as NASBA (nucleic acid sequence-based amplification) is incorporated[90] (Fig. 1.9B). In NASBA, reverse transcription takes place with the binding of target RNA template and reverse primer to generate a DNA/RNA duplex. Then the RNase H degrades the RNA temple to release the ssDNA. A forward primer containing the T7 promoter initiates the formation of dsDNA. Then the RNA polymerase generates numerous copies of new target RNA templates. Those new RNA templates can react with toehold switches and can also serve as new templates for further amplification. To release the RNA template from virus capsid, a simple heating step is needed. After diluted with water and heated up to 95˚C for 2 min, sufficient RNA templates can be released for running NASBA (Fig. 1.9C). With this technique, Zika samples with clinical level virus concentrations were successfully tested.

Figures

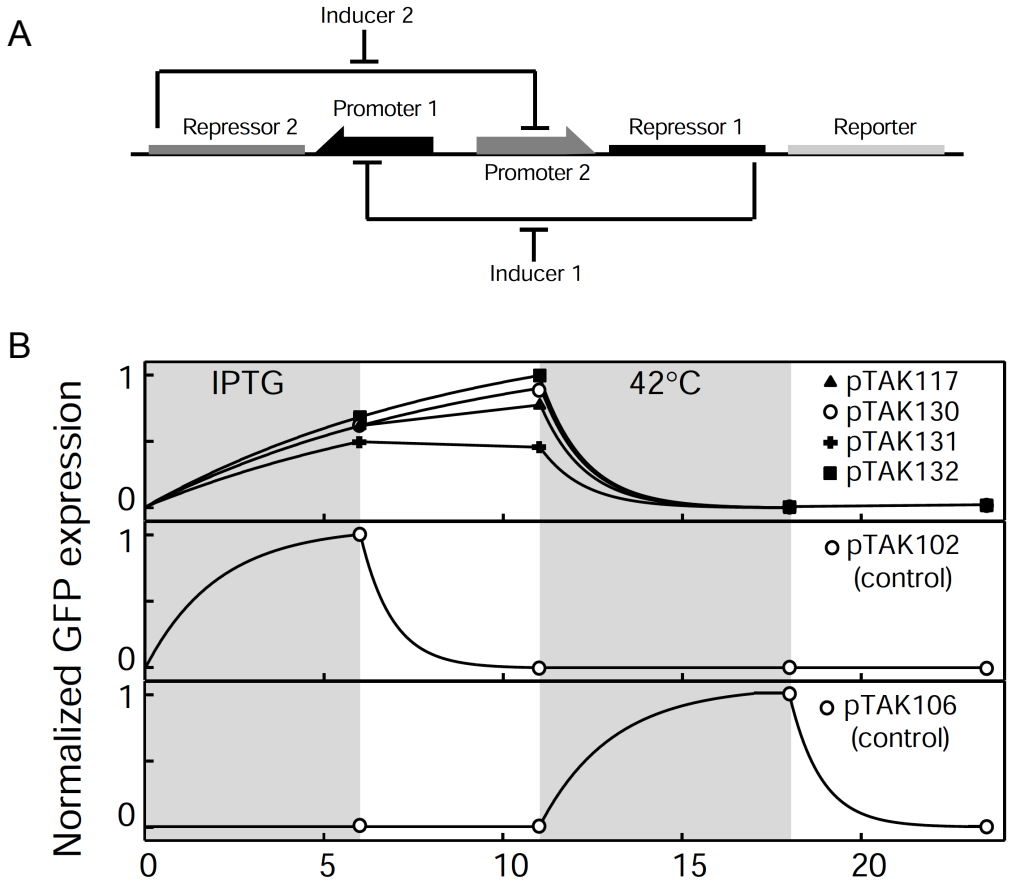


Figure 1.1 Toggle switch design and experimental results.

(A) Design of toggle switch. **(B)** Test of toggle switch in E. coli. The grey shading indicates periods of chemical or thermal induction.

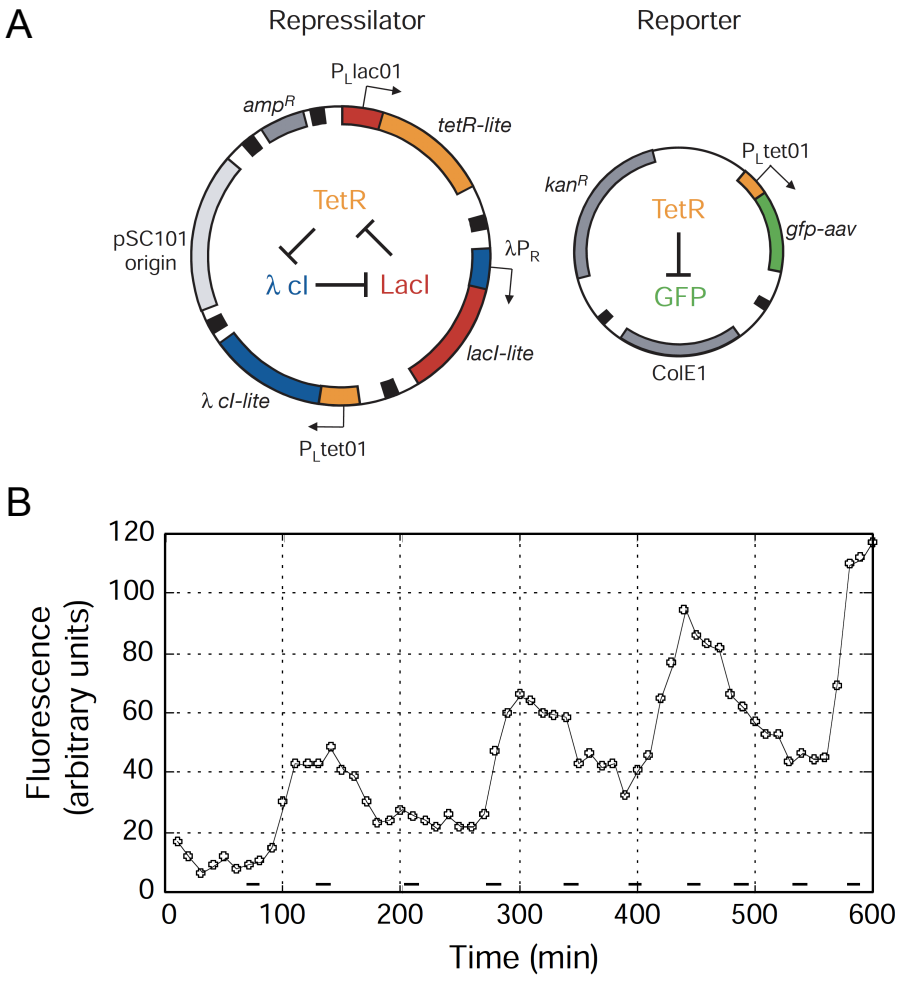


Figure 1.2 Oscillator design and experimental results.

(A) Plasmids construction of oscillator systems. **(B)** Time course GFP fluorescence from cells.

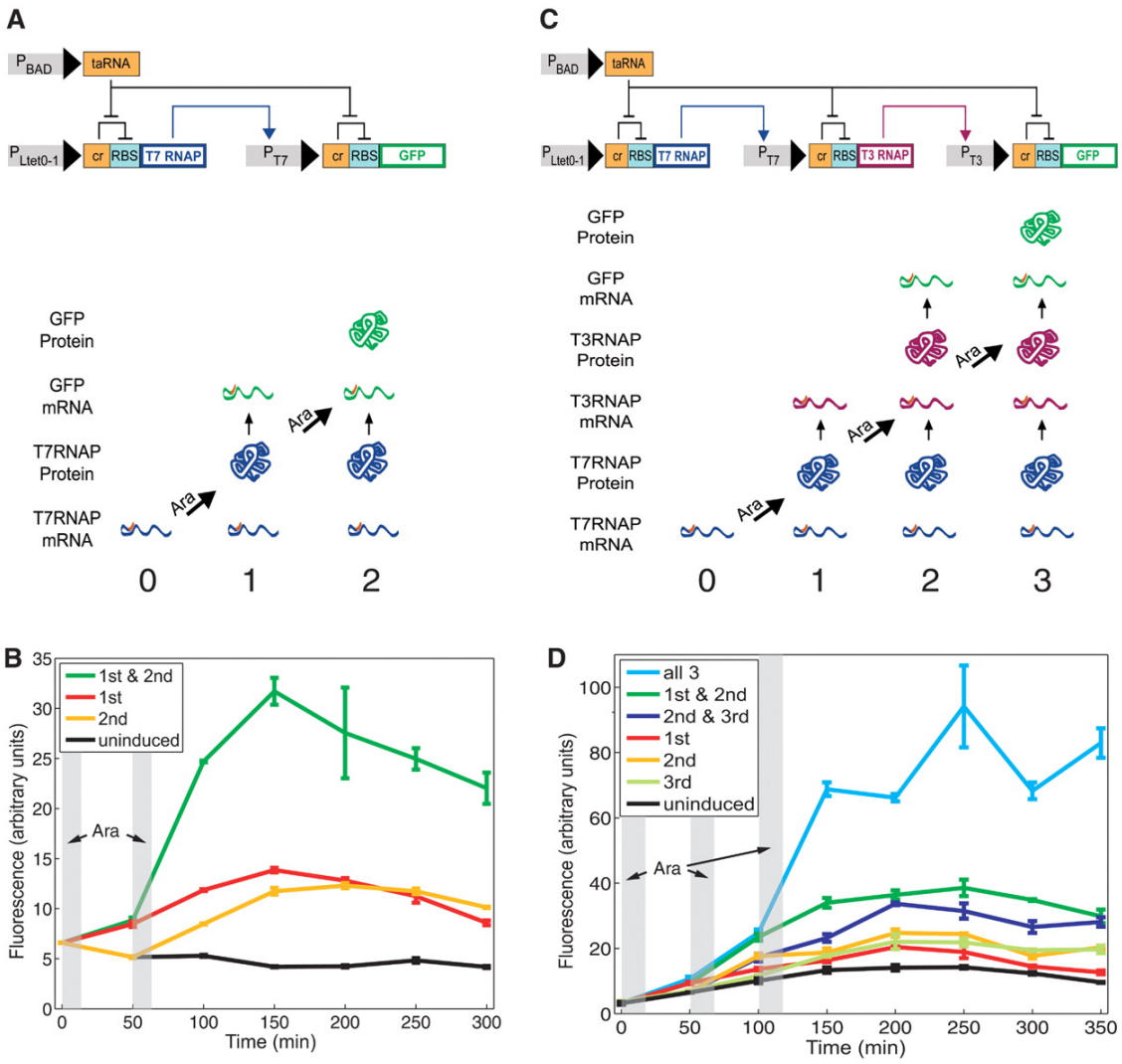


Figure 1.3 Counter designs and experimental results.

(A) The RTC two-counter design and expression profiles after arabinose pulses.

(B) Mean fluorescence of three replicates of two-counter systems via flow cytometry.

(C) The RTC three-counter design and expression profiles after arabinose pulses.

(D) Mean fluorescence of three replicates of three-counter systems via flow cytometry

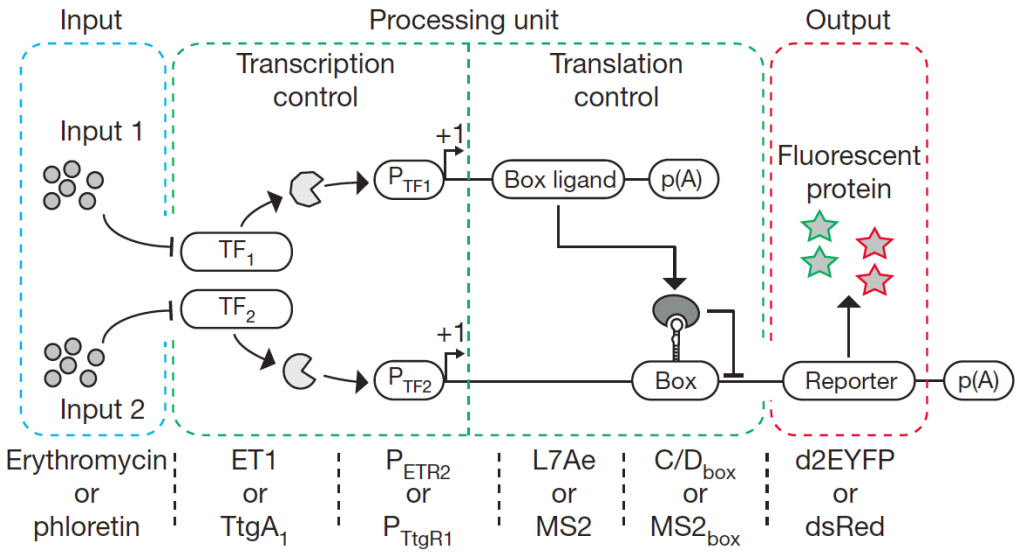


Figure 1.4 Genetic switchboard of the biocomputer circuitry.

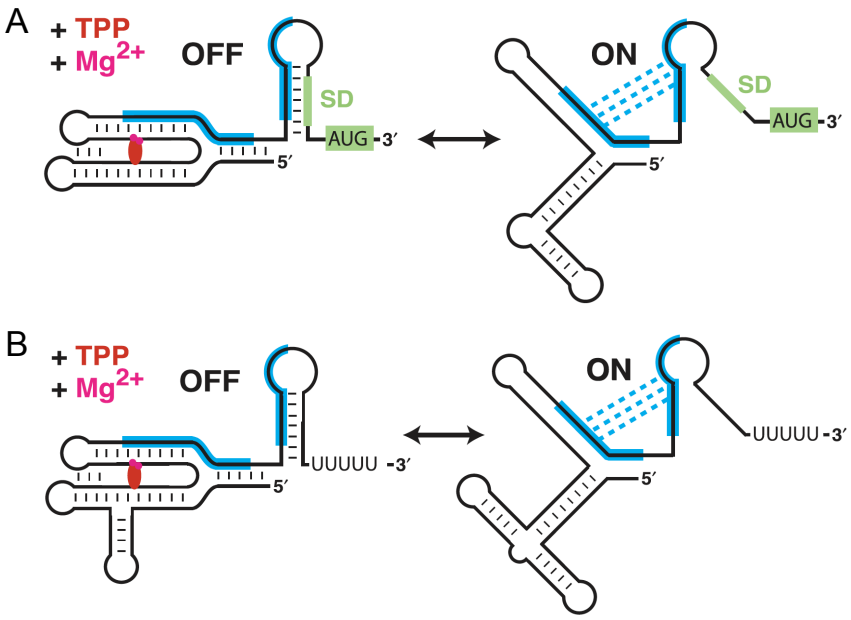


Figure 1.5 Schematic of the gene regulation by TPP-binding riboswitch[44].
(A) Regulation of translation initiation. **(B)** Regulation of transcription termination. Complementary sequences are highlighted in blue, and their interactions are labeled with blue dashed lines. SD sequence and start codon are labeled with green. TPP and Mg2+ are marked with red and magenta, respectively.

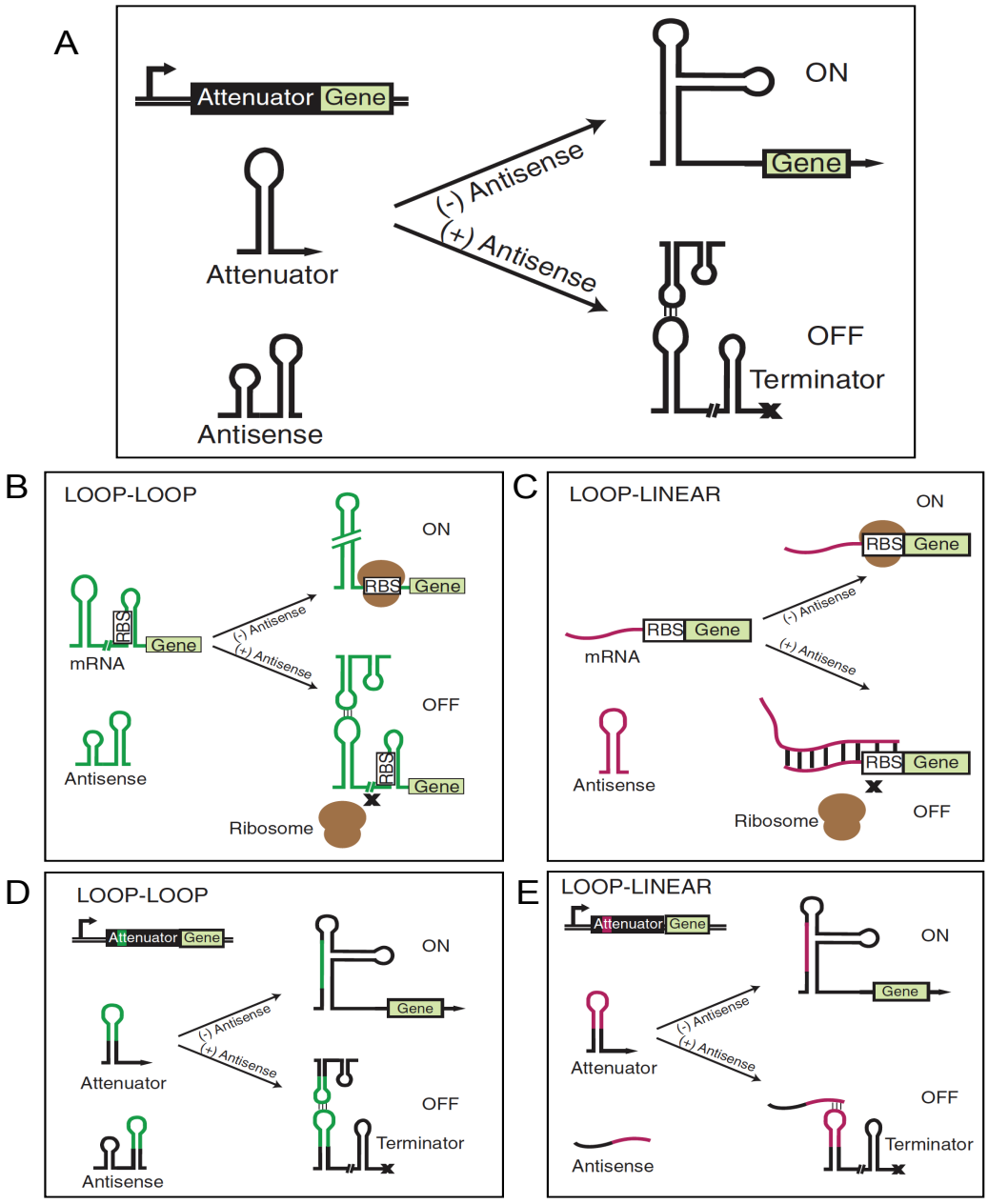


Figure 1.6 Gene regulation by natural and engineered antisense-RNAs[82].

(A) A transcriptional attenuator located at the 5' non-coding region of an mRNA can regulate gene expression at transcriptional level by regulating the formation of transcription terminators[80]. **(B, C)** Antisense RNA translational regulate gene expression at translational level by regulating the accessibility of RBS to ribosomes, with either a loop–loop interaction (B), or a loop–linear interaction (C). **(D, E)** Engineered transcription attenuators.

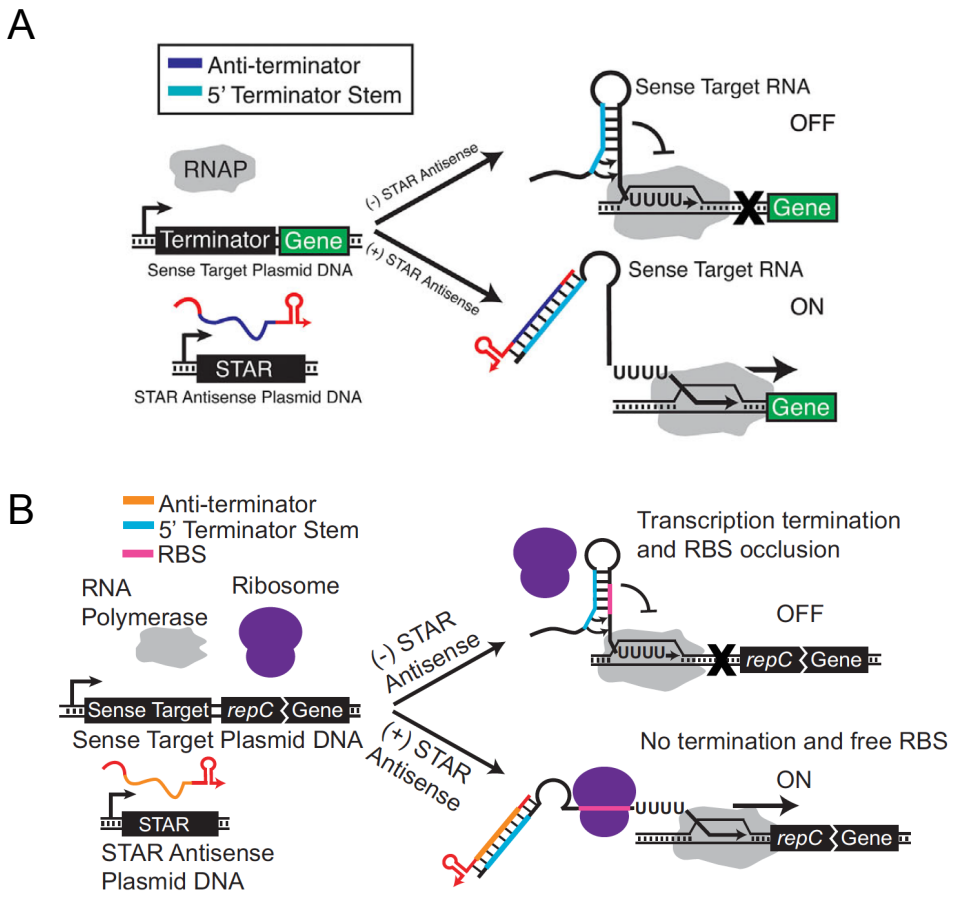


Figure 1.7 Small Transcription Activating RNAs (STARs)[49] and dual-level activator[83].
(A) Mechanism of STARs. The terminator formed during transcription in the mRNA will stop the transcription process. When the STAR disrupted the terminator structure during transcription, full length mRNA can be successfully transcribed, so that gene expression can be turn on.
(B) Mechanism of the transcriptional and translational dual-activator. Without STAR, terminator will stop mRNA transcription once formed, and the RBS domain will also be blocked in the formed hairpin structure. When STAR binds to its targeting domain, the terminator cannot form during transcription, and the RBS domain will also stay in a linear structure, which is accessible for ribosome binding to translate downstream proteins.

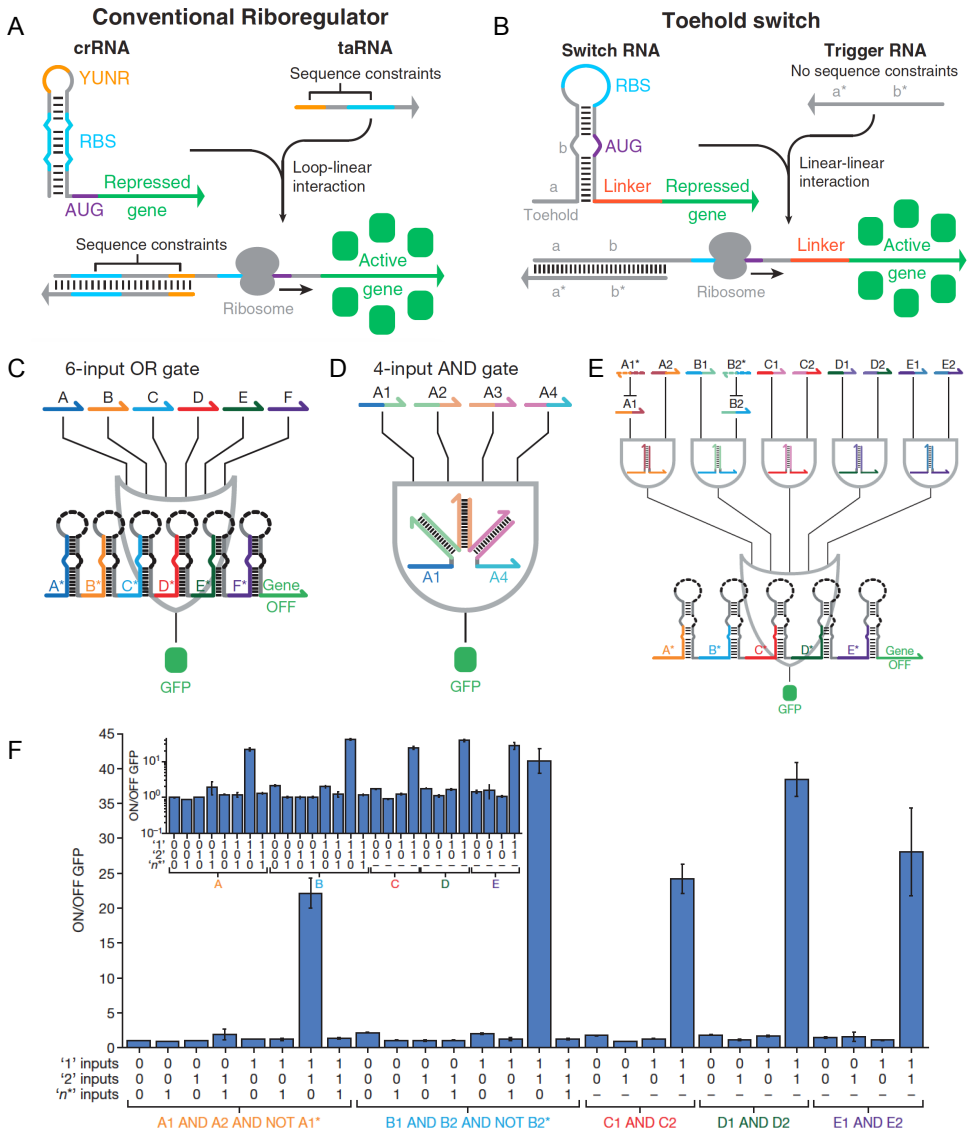


Figure 1.8 Toehold switch[84] and derived ribocomputing designs[85].

(A and B) Design schematics of conventional riboregulators and toehold switches.

(C) Schematic of the six-input OR gate RNA with six sensor modules.

(D) Schematic of the input RNA interaction used in a four-input AND gate.

(E) Schematic of the 12-input DNF expression evaluated in E. coli

(F) ON/OFF GFP from the DNF circuit under 28 different input RNA combinations; Inset, ON/OFF GFP on a logarithmic scale.

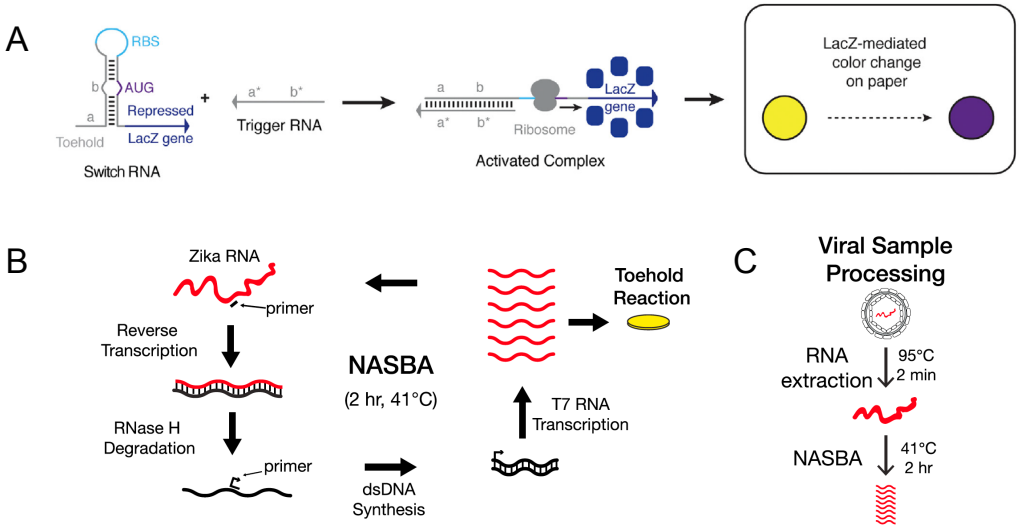


Figure 1.9 Modified toehold switch with colorimetric output[87] and NASBA reaction[89].

(A) Schematic of the LacZ-expressing toehold switches used to generate colorimetric outputs.

(B) Schematic of NASBA (nucleic acid sequence based amplification)-mediated RNA amplification.

(C) Pre-heating step of clinical samples.

CHAPTER 2

LOW-COST DETECTION OF NOROVIRUS USING PAPER-BASED CELL-FREE SYSTEMS AND SYNBODY-BASED VIRAL ENRICHMENT

2.1 Introduction

Noroviruses are the leading cause of human gastroenteritis[91] and globally are estimated to exact $60 billion in societal costs each year[92]. These viruses are extremely contagious[93] and can persist in the environment on contaminated surfaces, causing frequent outbreaks in closed settings such as long-term health care facilities, hospitals, schools, and cruise ships[91]. Noroviruses are also the primary cause of foodborne illness[94], and thus impose substantial risks to the food industry. Although norovirus infections are often self-limiting with symptoms that persist for only 48 hours in healthy individuals, the virus can lead to severe symptoms and prolonged illnesses in young children and the elderly[91]. In the United States alone, norovirus infections are estimated to impose an annual burden of 400,000 emergency department visits, 1 million pediatric medical care visits, and 19-21 million total illnesses[95,96]. In developing countries, the virus is associated with poorer health outcomes and is estimated to cause over 200,000 deaths annually[97,98].

Methods to accurately diagnose norovirus infections are thus essential to contain the spread of outbreaks and to ensure that patients receive optimal treatment. Noroviruses are non-enveloped, single-stranded, positive-sense RNA viruses and they exhibit substantial genetic variability, which can complicate efforts to develop effective diagnostics. Over 30 norovirus genotypes are known to cause acute gastroenteritis in humans. Of those genotypes, the GII.4 virus is the most commonly reported[99,100], and newly evolved strains in this genotype have emerged to cause global pandemics roughly every two to three years since the mid-1990s[91,93]. At present, the most commonly used methods to detect norovirus are immunochromatographic lateral flow assays and PCR-based tests[91]. Immunochromatographic lateral flow assays employ antibodies that recognize viral surface proteins and can be advantageous since they do not require specialized equipment and provide test results in 15 minutes. However, these tests provide limited sensitivity[101,102] and their results can be strongly genotype dependent[91]. Real-time

quantitative reverse transcriptase PCR (qRT-PCR) is currently the gold standard for detection of norovirus[91]. qRT-PCR assays can be targeted to conserved regions of the norovirus genome and they provide high specificity and sensitivity. These assays, however, require expensive thermal cycling equipment and are typically run in centralized laboratories[103]. Shipment of samples can delay test results and specialized laboratory equipment is often not available in developing countries or in remote settings, such as ships at sea, where outbreaks frequently occur. Highly automated commercial instruments such as the Cepheid GeneXpert have been developed for decentralized use[103]. However, these instruments are expensive. Even with negotiated prices for low- and middle-income countries, the GeneXpert instrument, for instance, costs $17,000 and has cartridges available at a concessional price of $9.98. These factors lead to an overall per test cost of $14.93 once labor, consumable, and other costs are included[104]. In the absence of discounts, GeneXpert costs rise substantially to $30.26-$155.44 per test depending on the country[105].

In response to these limitations, researchers have implemented nucleic acid tests for norovirus that employ isothermal amplification methods rather than conventional PCR[106-116] to obviate the need for expensive equipment and facilitate decentralized assays. Use of isothermal amplification methods, such as nucleic acid sequence-based amplification (NASBA)[117], loop-mediated isothermal amplification (LAMP)[118], and recombinase polymerase amplification (RPA)[119], avoids the need for expensive thermal cyclers, and it has enabled noroviruses to be detected in fecal samples[106-108,111-114,116], surface water[109], and in oysters[110,115]. However, these tests have required additional equipment to read out assay results via fluorescence or electrochemiluminescence[106,108,109,112,114,116], have the potential for false positives if the isothermal amplification is non-specific[107,110,111,113,115], or rely on expensive reagents such as non-canonical or fluorescent DNA bases[108,109,112,114,116]. We have recently developed an alternative strategy for nucleic acid detection that employs transcription-translation cell-free reactions embedded on paper substrates[87]. These systems employ freeze-drying to preserve the activity of the cell-free systems for over a year at room temperature, and they can be reactivated using water to enable synthetic biology tools to be deployed outside the lab. In previous work, these paper-based

systems successfully detected the Zika virus when coupled with isothermal NASBA reactions[89] and employed programmable riboregulators called toehold switches[84] to directly verify the sequence of the amplified RNA and produce the reporter enzyme ß-galactosidase (lacZ). This Zika assay was low cost at $1 to read out each test, required only inexpensive equipment, and provided results that could be detected directly by eye through cleavage of a chromogenic lacZ substrate.

Here we report the application of the paper-based cell-free platform for the detection of the prevalent GII.4 Sydney norovirus genotype (Fig. 2.1). Beginning from fecal samples or a dilute solution containing the virus, the assay employs biotin-labelled synthetic peptide affinity ligands known as synbodies to capture norovirus particles and concentrate them using streptavidin-coated magnetic beads. A brief heating step is used to release the norovirus RNA, and either NASBA or reverse transcriptase RPA (RT-RPA) is employed to amplify the viral RNA. The amplification products are then added to paper-based cell-free reactions where norovirus-specific toehold switches are used to verify their sequences and produce the lacZα peptide, which provides a visual reaction readout. We demonstrate that this assay enables detection of norovirus GII.4 Sydney from a stool sample down to concentrations of 270 aM without the use of a concentration step, and further show that synbody-based enrichment of the virus can lower the detection limit by 1000-fold to 270 zM. We also demonstrate that the use of α-complementation, in which the lacZα and lacZω peptides complement to form the active lacZ enzyme, can reduce the time to detection of the paper-based assay by up to 23 minutes or 43% compared to experiments employing the full-length lacZ as the toehold switch output. These results expand the range of sample types and viruses that can be analyzed using paper-based cell-free systems and provide new strategies to improve the sensitivity and reduce the time of these inexpensive diagnostic assays.

2.2 Materials and methods

2.2.1 Norovirus samples and bacterial strains

Stool samples positive for the norovirus GII.4 Sydney genotype and the norovirus GI.2 genotype were generously provided by Jan Vinjé from the National Calicivirus Laboratory at the Centers for Disease Control and Prevention (CDC). *Escherichia coli* MG1655 (ATCC, 700926), methicillin-resistant *Staphylococcus aureus* MRSA252 (ATCC, BAA-1720), and *Bacillus subtilis* 168 (ATCC, 23857) were used for assay cross-reactivity experiments. For these experiments, RNA from the bacteria was extracted using a Quick-RNA Fungal/Bacterial Miniprep Kit (Zymo Research) following the manufacturer's instructions. To obtain purified viral RNA for cross-reactivity experiments, 5 µL of GII.4, GI.2, and GI.6 positive stool samples were suspended in 140 µL RNase-free water. The viral RNA was extracted by using QIAamp DSP Viral RNA Mini Kit (Qiagen, U.S.A.) according to the manufacturer's instructions. RNAs were eluted with 50 µL RNase-free water and stored at -80°C. *E. coli* DH5α (ThermoFisher Scientific) was used for cloning of toehold switch plasmids.

2.2.2 *In silico* selection of toehold switch designs

An updated version of the selection algorithm described previously[84] was used to identify toehold switches for detection of norovirus RNA. The algorithm facilitated selection six promising designs from a set of over 100 candidate toehold switches generated from each norovirus target RNA. Candidate devices were designed to bind to a 36-nt continuous region of the norovirus target RNA. Putative toehold switches were generated at 1-nt increments along the norovirus target RNA and multiple ensemble defect levels were computed for each sensor based on its deviation from the ideal secondary structure of the toehold switch. Ensemble defects were calculated for the toehold switch 5' end through to the 3' end of the hairpin ($d_{\text{min\_sensor}}$), the toehold domain of the toehold switch (d_{toehold}), the binding site of the toehold switch within the target RNA ($d_{\text{binding\_site}}$), and the toehold switch region starting with the base immediately 3' of the target RNA binding site and extending 31 nts beyond the last base on the 3' end of the hairpin ($d_{\text{active\_sensor}}$).

The parameter d_{active_sensor} was intended to provide a measure of any secondary structures in the activated toehold switch that could interfere with translation after binding to the target RNA.

In addition to ensemble defects, the equilibrium fraction *f* of target/toehold switch complexes in a system with equimolar concentrations of target and toehold switch RNAs was calculated as a measure of the affinity of the two RNAs. In practice, this parameter was almost always equal to 1. Designs that produced in-frame stop codons in the output gene were eliminated from further consideration. Each of the parameters was then normalized such that their maximum value across the set of putative designs for a given target RNA was equal to 1. These normalized parameters, designated by an overscore, were then inserted into a scoring function *s*:

$$s = 5\bar{d}_{toehold} + 4\bar{d}_{active\_sensor} + 2\bar{d}_{min\_sensor} + 2\bar{d}_{binding\_site} + (1 - f)$$

Toehold switches displaying the lowest values of *s* and screened to have *f* > 0.9 were selected for experimental testing. The weighting coefficients used in the scoring function were determined empirically based on testing of earlier toehold switch mRNA sensor designs[84,87].

2.2.3 Toehold switch plasmid construction

Plasmids and DNA templates for transcription were constructed using conventional molecular biology techniques. Synthetic DNA (Integrated DNA Technologies) encoding the norovirus-specific toehold switch sensors was amplified by PCR and inserted into plasmids using Gibson assembly[120] with 30-bp overlap regions as described previously[121]. The sequences of the plasmids were confirmed using Sanger sequencing (DNASU Sequencing Core, Tempe). This table lists the source template amplified by each primer pair and indicates what plasmid was produced following Gibson assembly of the resulting PCR products.

2.2.4 Preparation of paper-based cell-free systems

Cell-free transcription-translation systems (NEB, PURExpress) were prepared for freeze-drying with the following components by volume: cell-free solution A, 40%; cell-free solution B, 30%; RNase Inhibitor (Roche, 03335402001, distributed by MilliporeSigma), 2%; chlorophenol red-b-D-galactopyranoside (Roche, 10884308001, distributed by MilliporeSigma, 24 mg/ml), 2.5%;

with the remaining volume reserved for toehold switch DNA, water, and lacZω peptide added to a final concentration of 2 μM. When testing the toehold switches expressed from a plasmid, the plasmid DNA was added to the cell-free reaction mix to a final concentration of 30 ng/μL. When testing toehold switches expressed from linear DNA, the DNA was added to the cell-free reaction mix to a final concentration of 33 nM.

Filter paper (Whatman, 1442-042) for housing the cell-free reactions was first blocked with 5% bovine serum albumin (BSA) overnight. After blocking, the paper was washed three times in water for 5 to 10 minutes. The paper was then heated to 50°C for drying and cut into 2-mm diameter paper disks using a biopsy punch. The disks were transferred into 200-μL PCR tubes and 1.8 μL of the cell-free reaction mix was applied to each disk. PCR tubes containing the paper disks were then flash frozen in liquid nitrogen and transferred into a lyophilizer to dry overnight. Measurements were performed on the resulting paper disks two to four days after the freeze-drying process was completed. The paper disks remained active for at least a month of room-temperature storage using conditions described previously[87], with the systems stored under nitrogen, shielded from light, and in the presence of silica gel desiccation packages.

2.2.5 Screening of norovirus-specific toehold switches

Norovirus target RNA was produced using T7 RNA polymerase-based transcription (Epicenter, ASF3257) from linearized DNA templates. 1.8 μL of a 5 μM solution of the target RNA was applied to a paper disk containing the embedded cell-free system and DNA for the toehold switch. The progress of the cell-free reaction was then monitored in a plate reader (Biotek, H1MF) at 37°C in triplicate. The relative absorbance of the paper-based reactions at 575 nm wavelength or OD575 was calculated by taking the absorbance at 575 nm and subtracting from it the absorbance at 575 nm measured at the start of the reaction. This relative absorbance thus removes any absorbance contribution from the paper disk and the lacZ substrate chlorophenol red-b-D-galactopyranoside. The fold change in lacZ production rate was calculated by computing the rate of change in OD575 and dividing the rate obtained for the toehold switch in the presence of the target RNA by that obtained in the absence of the target RNA. The fold change in lacZ

production rate was measured after one hour of cell-free reaction for assessment of the toehold switches. The change in OD575 or ΔOD575 was calculated by taking the OD575 for the reaction with the toehold switch and the target RNA and subtracting from it the OD575 for the reaction of the toehold switch without the target RNA. ΔOD575 was computed after two hours of cell-free reaction. Errors in OD575 were determined from the standard deviation of triplicate measurements. Errors in fold change lacZ production rate and ΔOD575 were determined by adding the relative and absolute errors of OD575 in quadrature, respectively. Welch's unequal variances t-test was used to calculate p-values for plate reader detection experiments with $p < 0.05$ used as the cutoff to define a statistically significant result.

2.2.6 Isothermal amplification of norovirus RNA

For NASBA experiments, reaction buffer (Life Sciences, NECB-24; 33.5%), nucleotide mix (Life Sciences NECN-24; 16.5%), RNase inhibitor (Roche, 03335402001; 0.5%), 12.5 mM of each DNA primer (2%), nuclease free water (2.5%), and RNA amplicon (20%) were assembled at 4°C and incubated at 65°C for 2 min, followed by a 10-min incubation at 41°C. Enzyme Mix (Life Sciences NEC-1-24; 25%) was then added to the reaction (for a final volume of 5 μL), and the mixture was incubated at 41°C for 2 hr. The amplified product was then diluted 1:6 in water and applied to paper disks containing the cell-free system and DNA for the toehold switch.

RT-RPA experiments used the commercial TwistAmp Basic RT kit (TwistDx). Reactions were prepared by combining 10 μM forward primer (4.8%), 10 μM reverse primer (4.8%), rehydration buffer, RNase Inhibitor (Roche, 03335402001; 4.4%), and RNA amplicon (22%) at room temperature and transferring the mixture to the freeze-dried reaction pellet. After mixing, 2.5 μL of 280 mM magnesium acetate (5%) was added to start the reaction and it was incubated at 41˚C for 5-7 minutes. The reaction tube was then inverted vigorously 8-10 times, spun down briefly, and returned to incubation at 41°C for 2 hr. The amplified product was then diluted 1:6 in water and applied to paper disks containing the cell-free system and DNA for the toehold switch.

For determination of assay detection limits, NASBA and RPA reactions were run in triplicate for each concentration of the target RNA or virus and applied to the paper-based toehold switch reactions as described above.

2.2.7 Synbody-based virus enrichment

A 30-µL volume of MyOne Streptavidin C1 streptavidin-coated magnetic beads (Life Technologies, U.S.A.), corresponding to 2.1×10^8 to 3.6×10^8 total beads, was added to Protein LowBind tubes (Eppendorf, U.S.A.). The bead storage solution was removed and the beads were washed three times with 1 mL of PBST (0.05% Tween 20 in 1x phosphate-buffered saline). The beads were then blocked with 3% BSA in PBST overnight at 4°C. The following day, the beads were suspended in fresh 3% BSA in PBST and blocked for an additional 2 hours. The beads were then washed three times with PBST and suspended in 30 µL of 1x PBS (phosphate-buffered saline) to yield a final suspension of blocked magnetic beads.

A dilution series of virus particles ranging from $1:10^3$ to $1:10^7$ was prepared by first taking a 1-µL aliquot of a norovirus GII.4 Sydney positive stool sample, and diluting it into 1 mL of PBS. The resulting $1:10^3$ sample was serially diluted by factors of ten into PBS to generate the rest of the dilution series. Biotin-labelled synbody[122] ASU1052 was then added to a concentration of 1 µM into each diluted sample and incubated with shaking for 1 hour at room temperature. The solutions were then added to the blocked streptavidin-coated magnetic beads and shaken for an additional 15 minutes at room temperature. The beads were washed three times with PBST and one time with PBS and then suspended with 50 µL water. The beads were incubated for 2 min at 95°C to release the viral RNA for analysis. 50 µL of each stool dilution was also incubated for 2 minutes at 95°C and used for comparison.

For cross-reactivity testing and tests of the assay against the GII.6 genotype, 1 µL of GII.4, GII.6, and GI.6 positive stool samples, as well as a norovirus-negative stool sample, were diluted into 1 ml of PBS and followed by the synbody enrichment procedure described above.

2.3 Results and discussion

2.3.1 Design of toehold switches for norovirus GII detection

We first identified conserved sequence regions of the norovirus GII genome suitable for isothermal amplification and toehold-switch-based detection. Over 400 norovirus GII complete and partial genome sequences were downloaded from the NCBI database and aligned. A 200-nt target sequence that was highly conserved across the norovirus GII genomes was identified for subsequent amplification and detection experiments. This conserved sequence ran from the C-terminal region of the viral RNA-dependent RNA polymerase through to the N-terminal region of VP1, the major capsid protein.

Toehold switches for detection of the target sequence were then generated based on an updated design first applied to the detection of the Zika virus. The updated toehold switch design previously provided lower leakage compared to earlier toehold switches[89] and was originally developed for evaluating AND logic expressions in *E. coli*[85]. As illustrated in Figure 2.2A, binding of a cognate target RNA to the updated toehold switch unwinds the lower half of the switch RNA hairpin and leaves the conserved upper stem-loop intact. This upper stem-loop is sufficiently weak to expose the ribosomal binding site (RBS) to enable translation to occur[85]. Unlike earlier toehold switch mRNA sensors, the updated systems do not employ an RNA refolding domain downstream of the start codon, which could hamper translation of the output gene.

Based on the modified operating mechanism of the toehold switches, we implemented an updated design selection algorithm to identify the toehold switches most likely to be effective at detecting the target RNA. This algorithm modelled the interaction of a series of toehold switches designed to bind along the target RNA in 1-nt increments using the NUPACK software package[123,124]. Ensemble defect levels and the affinity of the toehold switch for the target RNA were used to select designs most likely to perform well. Since the target RNA can be transcribed in either the sense or antisense direction following amplification, the top six toehold switches for the sense and antisense target RNAs were selected for experimental testing.

2.3.2 Faster RNA detection with toehold switches using α-complementation of lacZ

In previous work using paper-based cell-free systems, the lacZ enzyme has been used as the output gene for the toehold switch to produce a visible test result through cleavage of a chromogenic substrate[87,89]. LacZ, however, at 3.1 kb in length is a relatively long reporter gene compared to alternatives such as GFP (0.75 kb) and mCherry (0.72 kb), which leads to several drawbacks. In particular, the longer length of lacZ means that a greater fraction of the cell-free system resources is consumed during transcription and translation, which weakens the output from the assay, and longer times are required for the protein to be synthesized and fold, which increases the time required for the test.

In response to the above limitations, we investigated using α-complementation of lacZ to decrease assay times and strengthen output from the cell-free transcription-translation reactions. Alpha-complementation is a widely applied technique often used for screening cloning vectors. It works by dividing the lacZ enzyme into two peptides termed α and ω (Fig. 2.2B). The lacZ α-peptide (lacZα) consists of the first 50 to 59 residues from the N terminus of lacZ and the ω-peptide (lacZω) comprises the remaining ~970 lacZ residues. The complete lacZ must form a tetramer before it becomes catalytically active; however, lacZω cannot form a tetramer on its own as it lacks residues critical for assembly. As a result, both lacZα and lacZω must be expressed before complementation occurs and an active lacZ tetramer can assemble.

We thus implemented toehold switches that used lacZα as the output protein and added the much larger lacZω peptide as a pre-synthesized component to the paper-based cell-free reactions. Since lacZα is encoded in 180 bp, which is only ~6% of the length of the full lacZ gene, translation of each lacZα molecule should occur faster compared to lacZ and could in principle impose a substantially smaller burden on the cell-free system for each active lacZ tetramer formed. DNA encoding the norovirus-specific toehold switches was cloned into vectors upstream of the lacZα open reading frame. Following sequence confirmation, the resulting plasmids were tested in paper-based cell-free reactions supplemented with lacZω, and cleavage of the chromogenic substrate chlorophenol red-b-D-galactopyranoside was monitored using a plate reader. Figure 2.2C-F shows the results of these experiments with six toehold switches named

S1, S2, etc., for the sense orientation of the target RNA and six toehold switches named A1, A2, etc., for the antisense target orientation. All of the toehold switches were tested in parallel with reactions in which no target RNA was present. These experiments were then used to determine the fold change in the lacZ production rate and the ∆OD575 for each sensor. Three of the sense toehold switches provided ON/OFF ratios of approximately three or more (Fig. 2.2C) and displayed a change in absorbance at 575 nm (ΔOD575) of at least 0.4 (Fig. 2.2D), which can be discerned by eye. The toehold switches for the antisense target provided better performance overall with ON/OFF ratios up to 12.6-fold for A1 (Fig. 2.2E) and ΔOD575 up to 0.92 for A2 (Fig. 2.2F). Although the *in silico* selection algorithm successfully generated functional toehold switches for the two norovirus targets, we only detected appreciable correlations between the scoring function and the toehold switches for the antisense target. The sense target devices showed no correlations with the scoring function. Analysis of the experimental data indicates that other combinations of ensemble defect parameters coupled with different weighting factors can provide more accurate predictions of device performance.

To determine the effect of α-complementation on detection speed, we took one of the better performing toehold switches, A2, and inserted it into a plasmid upstream of the full lacZ open reading frame. PCR was then used to amplify linear DNA fragments from both lacZα and full-length lacZ plasmids and equal concentrations of the two DNA products were tested in paper-based cell-free reactions in the presence of the norovirus target RNA. We observed a substantial increase in the speed of the colorimetric reaction for the lacZα systems compared to full-length lacZ (Fig. 2.2G). Applying OD575 = 0.4 as the detection threshold, the lacZα reporter reached a positive result in 33 minutes compared to 56 minutes for the complete lacZ, which corresponds to a 40% reduction in detection time. Since both reactions reach saturation and completely cleave the substrate within the two-hour measurement shown in Figure 2.2G, we attribute the increased speed of the reaction in these conditions to the faster folding time of lacZα compared to lacZ, rather than to any decrease in the burden on the cell-free reaction caused by the shorter reporter protein.

2.3.3 Isothermal amplification using NASBA and RT-RPA

Since the concentrations of norovirus in stool samples from symptomatic patients range from ~30 aM to ~3 pM,[125] the toehold switches cannot be efficiently activated by viral nucleic acids without an amplification step. We investigated the NASBA and RT-RPA isothermal amplification techniques to determine which provided the lowest limit of detection against the norovirus GII.4 target RNA. The six toehold switches providing the highest ON/OFF ratios were selected for testing with amplified RNA. Since each sensor targeted different regions within the conserved target sequence, we evaluated different amplification primers for each sensor. One primer from each pair contained a 5' T7 promoter sequence so that the resulting amplicon could be transcribed into RNA for optimal detection using the corresponding toehold switch.

Toehold switches S2 and S6 provided the lowest detection limits in the amplification tests. Two-hour amplification reactions were run with synthetic norovirus GII.4 target RNAs ranging in concentration from 220 fM to 0.2 aM. The amplified products were then diluted seven-fold and applied to the toehold switch reactions. For the RT-RPA reactions, both S2 and S6 toehold switches could detect down to 22 fM of the norovirus RNA with colorimetric outputs that could be readily discerned by eye (Fig. 2.3A,B). Statistically significant concentrations as low as 2.2 fM could be detected from quantitative plate reader absorbance measurements for toehold switch S2 after 3 hours and toehold switch S6 after 1 hour.

NASBA tests provided improved detection limits compared to RPA. For toehold switch S6, we could discern concentrations down to 2 fM by eye within 2 hours and by plate reader within 1 hour (Fig. 2.3C). Although toehold switch S2 was not one of the very top performers in the initial screen (Fig. 2.2), it provided the lowest detection limit when coupled with NASBA. Experiments showed this sensor could detect down to 200 aM concentrations of the synthetic norovirus transcript (Fig. 2.3D). In addition, the sensor enabled detection by eye in 60 minutes at the 200 am detection limit as shown in Figure 2.3E and by plate reader in 28 minutes. A concentration of 200 aM corresponds to 600 copies of the RNA template in the 5 µL NASBA reaction.

2.3.4 Diagnostic validation with active norovirus

To validate the detection platform, we performed experiments with active norovirus samples and tested the assay for cross-reactivity against other potential pathogens. Following previous reports on norovirus[116,126] and our earlier work on the Zika virus[89], we first evaluated a simple method for extracting viral RNA from infected stool samples using a brief heating step. A norovirus GII.4 Sydney positive stool sample was diluted 1:50 in PBS and heated for two minutes at 95°C (Fig. 2.4A). The same procedure was applied to a stool sample not infected with the virus and two additional stool samples containing norovirus GI.2 and GI.6. These heated samples, along with comparison unheated samples and a water-only negative control, were both amplified by NASBA over 2 hours and applied to a paper-based reaction with toehold switch S2. The unheated samples all yielded minimal changes in toehold switch output compared to the negative control. The OD575 of the heated sample with norovirus GII.4 Sydney increased to 1.13, while the OD575 of the other heated samples remained below 0.25 (Fig. 2.4B). Thus, the simple heating method was effective at releasing RNA from norovirus particles and the assay was specific for norovirus GII.4 Sydney.

To further evaluate cross-reactivity, we extracted RNA from *E. coli*, *B. subtilis*, and a methicillin-resistant *S. aureus* (MRSA) strain and added the RNA at masses of 80.6 ng, 123.5 ng, and 100.8 ng, respectively, to the NASBA reaction. RNA was also extracted from stool samples containing norovirus GII.4 Sydney, GI.2, and GI.6 and added to the NASBA reaction at a concentration of approximately 20 fM. None of these samples of bacterial RNA nor the GI.2 and GI.6 norovirus genotypes were able to activate toehold switch S2 for visual detection. The system was strongly activated by norovirus GII.4 Sydney RNA (Fig. 2.4C).

2.3.5 Norovirus enrichment using a synbody-based magnetic bead technique

The ability to identify norovirus in dilute solutions or from large solution volumes is valuable for improving diagnostic sensitivity and for confirming complete decontamination of an area following an outbreak. For instance, dilute liquids, such as cleaning solutions from kitchen and bathroom surfaces, can be tested for residual virus following cleanup. To this end, we employed a synbody-based magnetic bead capture assay to concentrate norovirus from dilute

solutions (Fig. 2.5A). Synbodies are synthetic bivalent affinity ligands composed of two 15- to 20-mer peptides screened to bind to the surface of a protein of interest. Synbodies have affinities and specificities similar to antibodies[127,128]. Unlike antibodies, however, which often lose their affinity as norovirus strains evolve[91], synbodies have broad cross-affinity for multiple norovirus genotypes, which enables them to recognize a range of norovirus genotypes within both the GI and GII genogroups[122].

To capture and concentrate the virus, we took a stool sample positive for norovirus GII.4 Sydney at a concentration of 270 fM as determined by qRT-PCR and prepared a series of higher dilutions ranging from $1:10^3$ to $1:10^7$ in PBS. Biotin-labelled synbody ASU1052, which was previously validated against multiple norovirus strains[122], and streptavidin-coated magnetic beads were added sequentially to the diluted samples with shaking at room temperature for 75 minutes total. After magnetic capture and washing, the beads were suspended with 50 µL of water and heated to 95°C for 2 min to release the virus RNA. These virus samples, along with comparison ones heated but not subjected to synbody capture, were then amplified using NASBA and applied to paper-based cell-free systems containing toehold switch S2.

Figure 2.5B displays the absorbance change produced from the reactions after two hours with the two different sets of samples. For synbody-concentrated samples, norovirus could be detected by eye with dilution factors up to 10^5, which corresponds to a concentration of 2.7 aM. In contrast, none of the samples used directly and not subjected to concentration could be detected within two hours by eye. To further compare the two preparation methods, Figure 2.5C shows the absorbance change over time for several virus samples. The synbody-concentrated sample prepared from a 10^5 dilution crosses the eye-based detection threshold of OD575 = 0.4 in under two hours and provides a statistically significant positive signal in the plate reader after 66 minutes. The profile of the non-concentrated sample diluted 1000-fold nearly matches that of the synbody-concentrated sample diluted 10^6-fold over the full 4-hour measurement. Both samples cross the visual detection threshold after 3 hours and provide positive results from quantitative plate reader measurements in approximately 2 hours, which correspond to norovirus GII.4 Sydney detection limits of 270 aM and 270 zM for the non-concentrated and synbody-

concentrated samples, respectively. The synbody-based concentration technique thus enables a 1000-fold improvement in the detection limit of the norovirus assay.

To determine if the assay could also be applied to closely related norovirus genotypes, we also tested the systems against a stool sample with the norovirus GII.6 genotype. Virus particles were enriched using the ASU1052 synbodies and subject to NASBA using the primers optimized for GII.4 Sydney amplification. Unfortunately, these primers were not effective for this genotype. Primers modified to match the GII.6 genome, however, enabled successful amplification. Despite the presence of some mismatches between toehold switch S2 and its binding site on the GII.6 amplicon, a visible OD575 signal was observed from paper-based reactions within two hours (Fig. 2.4D). Thus, toehold switch S2 is capable to detect amplicons from both the norovirus GII.4 Sydney and GII.4 genotypes.

2.4 Conclusion

We have demonstrated a paper-based assay for detection of norovirus that does not require expensive thermal cycling equipment, provides test results that can be read directly by eye, and employs toehold switch riboregulators to eliminate false positives caused by non-specific amplification. The assay enables visual detection of norovirus down to a concentration of 270 aM from clinical stool samples containing live norovirus particles from the GII.4 Sydney genotype. The addition of a virus capture and concentration step using synbodies enables a further 1000-fold improvement in the sensitivity of the assay, allowing concentrations as low as 270 zM to be detected by eye after a three-hour paper-based reaction. This work also demonstrates that paper-based transcription-translation systems can remain active upon exposure to samples diluted from stool and confirms that RPA products can be successfully detected in the cell-free reactions, albeit with a higher detection limit than comparison NASBA products.

The norovirus assay provides significant improvements in sensitivity compared to our previously reported diagnostic assay for the Zika virus[89]. The Zika virus test provided a 1 fM detection limit against synthetic target RNAs and detected the virus from plasma at a

concentration of 2.8 fM. In contrast, the norovirus assay demonstrated a 5-fold lower detection limit of 200 aM against a synthetic target and was successfully applied to a stool sample with a 270 aM concentration of norovirus. Addition of the synbody concentration step thus yielded an overall 5000-fold improvement in the detection limit. The Zika virus is known to be present at very low levels in symptomatic patients, with serum concentrations ranging from 8 zM to 6.1 fM with an average of 160 aM[129]. These concentrations are 10- to 100-fold lower than those observed for patients with the related dengue and chikungunya viruses[130]. Accordingly, our synbody-based concentration methods could prove valuable for extending the existing Zika test to more carriers of the virus. While the Zika diagnostic was only applied to a plasma sample from a viremic rhesus macaque, we have also demonstrated in this work that the diagnostic platform can be used on human stool samples, which can be used to identify many other causes of acute gastrointestinal illness beyond norovirus.

Although our norovirus assay provides sufficient sensitivity for detection from clinical samples, at present it requires 3-6 hours of processing time to reach a test result, which is substantially longer than many other diagnostics that employ isothermal amplification. We expect that large reductions in assay time can be obtained by further optimization of the synbody-based enrichment technique, by designing toehold switches optimized for quicker and stronger output, and by implementing new reporter proteins with faster activation. Indeed, the substantial decrease in reaction time that we observed using α-complementation of lacZ suggests that there is ample room for improvement using alternative reporters. Moreover, use of faster amplification techniques such as RT-RPA with improved primers or strand-displacement amplification (SDA) could further decrease the time to detection for the technique. We also expect that toehold switch dynamic range against pathogen RNAs can be improved with continued refinement of *in silico* selection algorithms. In particular, screening experiments examining larger numbers of toehold switches against diverse target RNAs will be essential for generating *in silico* design scoring functions that are able to accurately predict their performance when deployed in cell-free transcription-translation systems.

The assay can also be improved by reducing its cost. In addition to the ~$1/test price of the paper-based component of the assay[87], the per test costs of NASBA, streptavidin-coated magnetic beads, and biotinylated synbodies are $2.25, $5.38, and $0.10, respectively. The total cost in materials for the assay is thus $8.73 and the overall assay requires approximately 35 minutes of hands on time. A previous study in South Africa to assess GeneXpert cartridge costs has reported an average lab technician salary of $9.07/hr,[104] which brings the total assay cost to $14.02 with labor included. Materials costs for this estimate are based on retail prices for the components. It is likely that the quantities of magnetic beads used in the assay can be reduced substantially with further refinement of the experimental procedures, and materials costs can decrease with purchases at larger scales. Even without optimization of the assay toward reduced price, the total cost per assay remains lower than the $14.93 calculated for GeneXpert cartridges in South Africa where concessional pricing is in effect[104]. Furthermore, our assay does not require large initial expenditures for purchasing expensive equipment.

The continual emergence of new variants of norovirus means that our paper-based assay will need to be updated as other strains replace GII.4 Sydney to ensure that false negatives do not occur. For instance, the GII.P17-GII.17 norovirus strain has recently become predominant in Asia[100] and immunochromatographic tests, which were developed for the GII.4 strain, have demonstrated 1000-fold poorer detection limits against the emergent strain[131]. To reduce the probability of false negatives, our assay employs a target sequence that is well conserved across different GII strains, including GII.P17 and GII.17. The toehold switch S2 sensor is predicted by NUPACK simulations to tolerate several mismatches in the target RNA, particularly within the toehold region, and still expose the RBS and start codon tzo enable translation of the reporter gene. This resiliency against sequence variations is evidenced by the ability of device S2 to activate against the GII.6 strain (Fig. 2.5D). In cases where there is larger sequence divergence, sensor mRNAs that employ multiple toehold switch hairpins upstream of a single output gene can be used to detect different norovirus strains or to compensate for locations with higher sequence variability to avoid false negatives. We have demonstrated that such OR logic systems can be used to detect six completely sequence-independent target RNAs using a single sensor mRNA in

E. coli[85]. We expect that similar approaches can be used in the paper-based reactions and prove more parsimonious with cell-free systems resources than other implementations employing multiple independent mRNAs. Like other nucleic acid tests that employ amplification, false negatives can also occur when the amplification primers do not have sufficient homology with the target amplicon. Such sequence variability can be addressed using primers with degenerate bases at positions known to have high probability of sequence divergence.

Despite these areas for improvement, the reasonably low cost of the assay and its reliance on only inexpensive equipment enables it to be implemented in decentralized contexts such as remote clinics or cruise ships with trained operators. Furthermore, coupling the validated molecular components of the assay with companion hardware for incubation and readout[89] or liquid handling[132] has the potential to substantially reduce operator training requirements and lead to more widespread deployment in the future. Lastly, the demonstrated ability of synbodies and toehold switches to bind to proteins and nucleic acids, respectively, from a variety of different pathogens[87-84,127,128] indicates that our combined concentration and detection approach can be successfully applied to a diverse range of infectious agents.

Figures

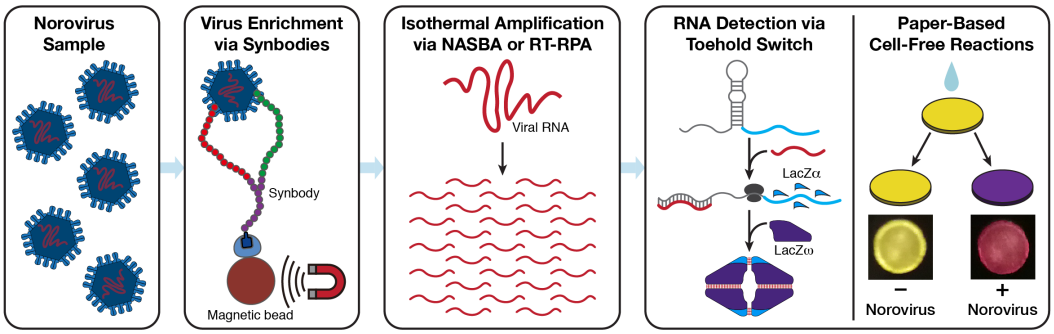


Figure 2.1 Overview of the norovirus detection assay using paper-based cell-free transcription-translation reactions. A norovirus sample is first enriched using synbodies and viral RNA amplified isothermally using nucleic acid sequence-based amplification (NASBA) or reverse transcriptase recombinase polymerase amplification (RT-RPA). The amplified nucleic acids are added to paper-based cell-free reactions where norovirus RNAs are detected by sequence-specific toehold switches. The toehold switches generate the lacZα peptide, which produces a purple-colored product after complementation with lacZω. Samples positive for norovirus can be identified by their purple color following the assay.

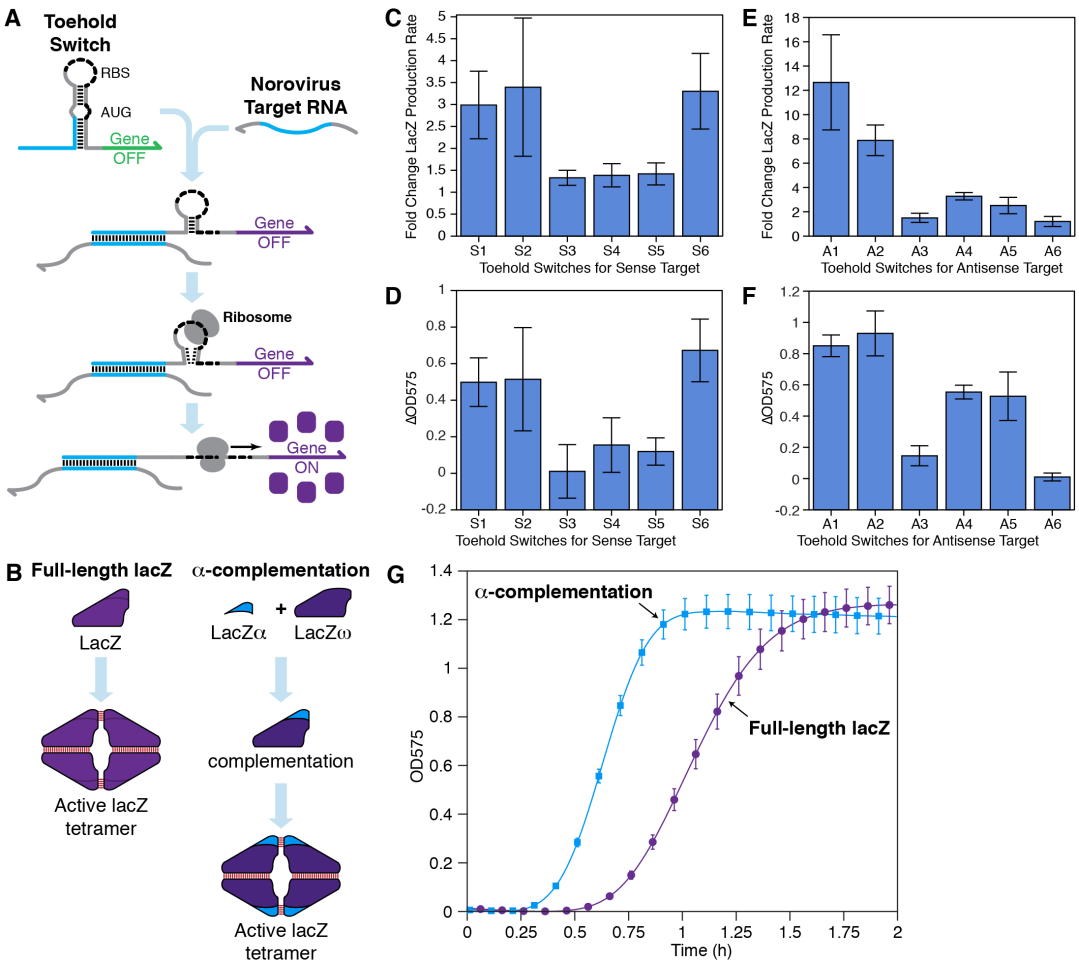


Figure 2.2 Detection of norovirus target RNA using toehold switches and α-complementation.

(A) Schematic of toehold switch operation in response to the target RNA. A weak stem containing the ribosomal binding site (RBS) is retained after target binding. This stem unwinds during binding of the ribosome to enable translation of the output gene.

(B) Enzymatically active lacZ tetramer formation occurs directly for full-length lacZ, while lacZα and lacZω must first assemble via α-complementation prior to tetramer formation.

(C and D) Measurements of the fold change in lacZ production rate (C) and ΔOD575 (D) of six toehold switches targeting the sense orientation of the norovirus target RNA.

(E and F) Measurements of the fold change in lacZ production rate (E) and ΔOD575 (F) of six toehold switches targeting the antisense orientation of the norovirus target RNA. Change in lacZ production rate was measured after 1 hour of cell-free reaction (C, E) and ΔOD575 was measured after 2 hours of cell-free reaction (D, F).

(G) OD575 for toehold switch A2 as function of cell-free reaction time when outputting full-length lacZ compared to lacZα in a reaction supplemented with pre-synthesized lacZω.

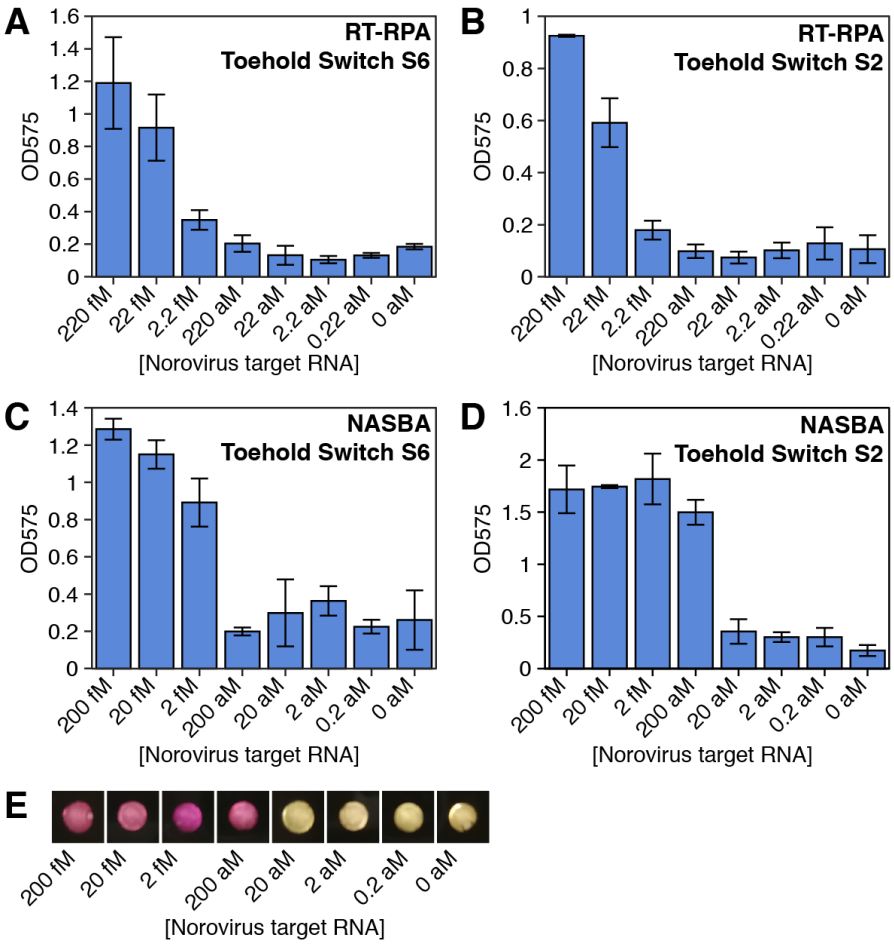


Figure 2.3 Detection limit measurements for synthetic norovirus GII.4 target RNAs subject to isothermal amplification and detection using toehold switches.

(A and B) OD575 after amplification using RT-RPA and detection using toehold switches S6 (A) and S2 (B) in two-hour cell-free reactions.

(C and D) OD575 after amplification using NASBA and detection using toehold switch S6 (C) and S2 (D) in two-hour cell-free reactions.

(E) Photographs of paper-based reactions using NASBA for amplification and toehold switch S2 for detection. Photographs were taken after 1 hour of the cell-free reactions.

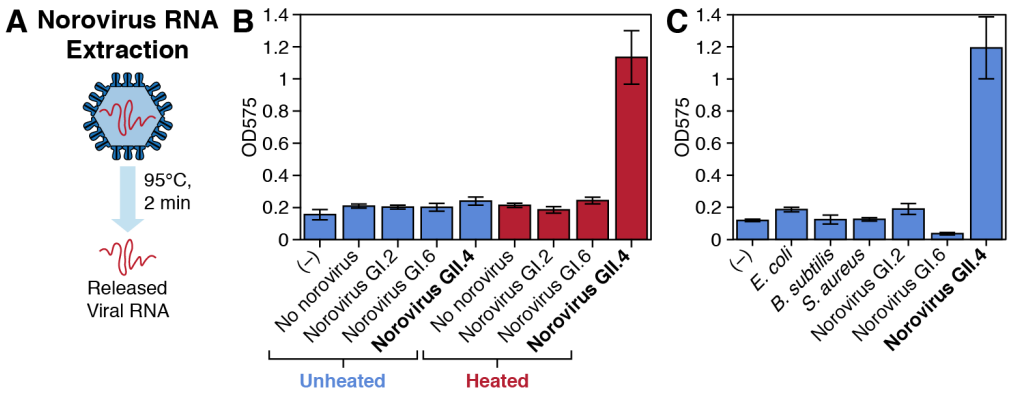


Figure 2.4 Detection of live norovirus GII.4 Sydney and cross-reactivity testing.

(A) Norovirus RNA was extracted by diluting a stool sample 1:50 into PBS and briefly heating to 95°C for 2 minutes.

(B) Measurement of OD575 after a two-hour paper-based reaction for a water-only negative control (–) and stool samples with and without norovirus particles before and after the brief heating treatment. All samples were subject to amplification via NASBA and detection with toehold switch S2. Only the heated norovirus GII.4 Sydney sample activates the toehold switch.

(C) Cross-reactivity testing of the assay against RNA from multiple bacteria, norovirus genotypes, and a water-only negative control. All samples were subject to NASBA and toehold switch S2 detection. OD575 was measured after two hours of the cell-free reaction.

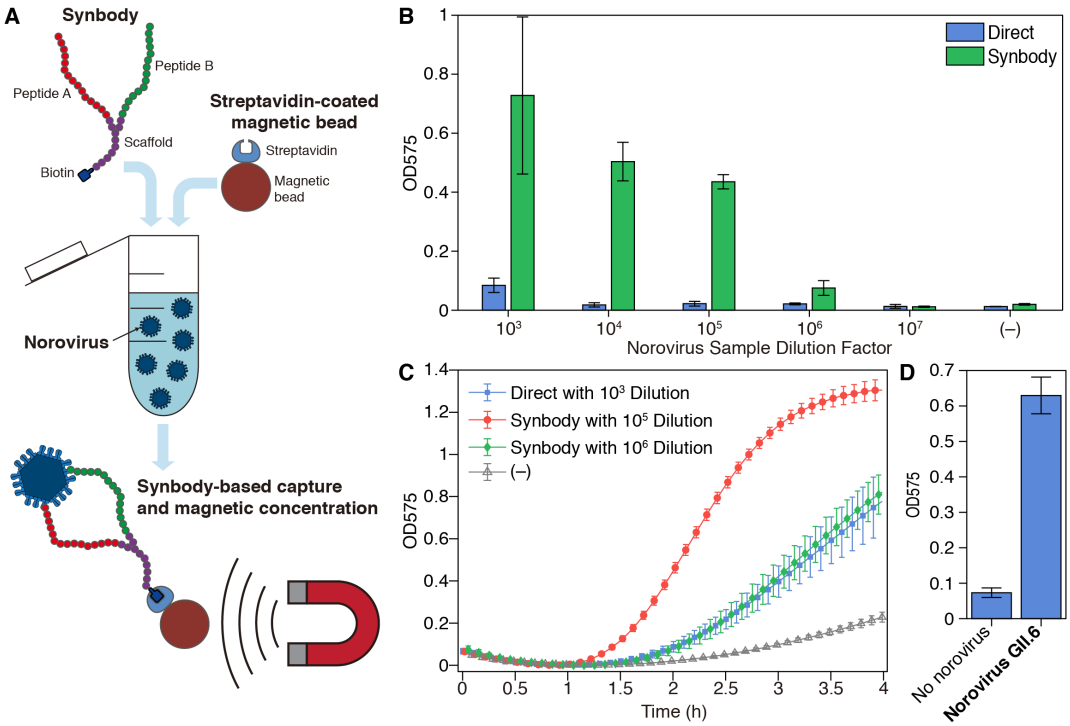


Figure 2.5 Implementation of a synbody-based capture and concentration method for norovirus detection.

(A) Illustration of the synbody enrichment technique. Biotin-labelled synbodies engineered to recognize diverse norovirus genotypes are used to bind to virus particles in a dilute solution and are in turn captured by streptavidin-coated magnetic beads. Magnetic capture enables concentration of the captured norovirus particles.

(B) Measurement of OD575 after two-hour cell-free reactions with toehold switch S2. Samples subject to synbody-based concentration and samples used directly without concentration were amplified by NASBA. The negative control (–) is a water-only sample.

(C) Time-course measurements of OD575 for synbody-concentrated samples compared to samples used directly. OD575 for a sample used directly after 1000-fold dilution is comparable to a concentrated sample initially diluted by 106-fold.

(D) Detection of norovirus GII.6 from a stool sample using toehold switch S2 and updated NASBA primers for the GII.6 genome. OD575 measurements were taken after two hours of the paper-based cell-free reaction and using a norovirus-negative stool sample as comparison.

CHAPTER 3

LOOP-INITIATED RNA ACTIVATORS: HIGH-PERFORMANCE TRANSLATIONAL REGULATION WITHOUT SEQUENCE CONSTRAINTS

3.1 Introduction

Synthetic biology unites biology and engineering[9,133] with a focus on constructing novel biomolecular components, complex synthetic networks, and unnatural metabolic pathways in living cells. To date, a number of biological devices have been constructed, including toggle switches[10], oscillators[11], molecular counters[134], logic gates[23,135-137], cell classifiers[138] and analog signal processors[139]. Yet efforts to create more complex devices have been hindered by the limited number of effective components that have high dynamic range and low crosstalk, especially when being put into the complex cellular environment[9].

RNA based components are an attractive means to construct more complex circuits since they can take advantage of predictable Watson-Crick base pairing and exploit the numerous RNA based gene regulation systems have been found in nature[140-142]. Indeed, researchers have created many engineered RNA riboregulators based on natural systems[143-148] that regulate transcription or translation upon detection of a target RNA. However, these systems have been limited in their dynamic range and their orthogonality. More recently, *de novo* designed riboregulators called toehold switches have been developed that can provide ON/OFF ratios over 600-fold and exhibit very low crosstalk levels across dozens of different devices[84]. Although these performance metrics represent substantial improvements over previous riboregulators, toehold switches do suffer from a number of important limitations. First, they exhibit noticeable signal leakage in vivo, which limits their use in applications requiring very stringent regulation of gene expression. Second, they impose restrictions on the sequence of the target RNA to be detected and the residues incorporated into the output gene.

We have recently developed a new type of synthetic riboregulator inspired by molecular beacons[149] that overcomes many of the limitations of toehold switches. These loop-initiated RNA activators (LIRAs) mechanism between a repressed mRNA containing a stem-loop and a single-stranded cognate target RNA. Testing of the LIRAs in E. coli has demonstrated that these

systems provide high dynamic range, low crosstalk and extremely low leakage in vivo. LIRAs can also be engineered to detect RNAs with arbitrary sequences, since no correlated domains with predefined sequences located in the trigger RNA binding region. By incorporating with paper-based diagnostic platform, LIRA realized the detection of viruses down to 20 aM, which is equal to 12 copies/μl.

3.2 Materials and methods

3.2.1 Strains and growth conditions

The following E. coli strains were used in this study: BL21 Star DE3 (F− ompT hsdSB (rB−mB−) gal dcm rne131 (DE3); Invitrogen), BL21 DE3 (F− ompT hsdSB (rB−mB−) gal dcm (DE3); Invitrogen), MG1655Pro (F− λ− ilvG- rfb-50 rph-1 SpR lacR tetR), and DH5α (endA1 recA1 gyrA96 thi-1 glnV44 relA1 hsdR17(rK−mK+) λ−; Invitrogen). All strains were grown in LB medium at 37 °C with appropriate antibiotics.

3.2.2 Plasmid construction

Plasmids were constructed by PCR and Gibson assembly techniques. Single-stranded DNAs for expressing LIRAs and target RNAs were purchased from Integrated DNA Technologies and amplified into double-stranded DNA via PCR. The amplified DNAs were then connected with plasmid backbones by 30-bp homology domains using Gibson assembly[120]. All Gibson assembly products were transformed in the E. coli DH5α strain and then sent out for sequence validation via Sanger sequencing. Backbones used for constructing the plasmids were amplified from the commercial vectors pET15b, pCOLADuet via PCR followed with Dpn1 treatment. The reporter protein for all plasmids is GFPmut3b with an ASV degradation tag.

3.2.3 Flow Cytometry Measurements and Analysis

Bacteria colonies transformed with LIRA and target RNA plasmids are inoculated in 1ml LB in triplicates with corresponding antibiotics. On the second day, 5ul overnight cultured medium is diluted by 100 folds in 495ul fresh LB with 60% kanamycin, 50% ampicillin and 50%

spectinomycin. After 80mins recover, IPTG is added into each well with a final concentration of 0.1mM. And flow cytometry measurement is performed after 3, 4 and 5hours inducement.

Flow cytometry was performed using a S1000 cell analyzer (Stratedigm) equipped with a high-throughput auto sampler (A600, Stratedigm). Before running measurement, cells were diluted by ~10 folds into phosphate buffered saline (PBS) in 384-well plates. Forward scatter (FSC) was used for trigger, and 40,000 individual cells were recorded. Cell populations were gated according to their FSC and side scatter (SSC) distributions as described previously. The GFP fluorescence signal outputs of these gated cells were used for following calculations. Error levels for the fluorescence measurements of ON state and OFF state cells were calculated from the SD of measurements from at least three biological replicates. The relative error levels for the ON/OFF fluorescence ratios were then determined by adding the relative errors of ON and OFF state fluorescence in quadrature.

3.2.4 Cell-Free Reactions

Cell-free transcription-translation systems (NEB, PURExpress) were prepared for freeze-drying according to following recipe: cell-free solution A, 40%; cell-free solution B, 30%; RNase Inhibitor (Roche, 03335402001, distributed by MilliporeSigma), 2%; chlorophenol red-b-D-galactopyranoside (Roche, 10884308001, distributed by MilliporeSigma, 24mg/ml), 2.5%; with the remaining volume reserved for LIRA DNA, water and lacZα peptide added to a final concentration of 2 μM. When testing LIRA expressed from a plasmid, the plasmid DNA was added to a final concentration of 30ng/μl to the cell-free reaction mix.

Filter paper (Whatman, 1442-042) for deposit and freeze-drying the cell-free system was first blocked with 5% bovine serum albumin (BSA) overnight. The paper was washed three times in water for 5 to 10 min after overnight blocking. The paper was transferred on a hot plate at 50˚C for drying and then cut into 2mm diameter paper disks with a biopsy punch. The disks were then transferred into 200μl PCR strips and 1.8μl of the above cell-free reaction mix was applied to each of them. Liquid nitrogen was used for freeze the PCR strips containing those paper devices. The frozen paper disks were dried overnight with a lyophilizer. Plate reader tests were carried out

on the freeze-dried paper disks 2-4 days later. With the systems stored in nitrogen environment, shielded from light and together with the silica gel desiccation packages as described previously, the paper disks remained active for at least a month under room temperature.

3.2.5 NASBA

NASBA experiments were carried out following the standard protocols: reaction buffer (Life Sciences, NECB-24; 33.5%), nucleotide mix (Life Sciences NECN-24; 16.5%), RNase inhibitor (Roche, 03335402001; 0.5%), 12.5 μM of each DNA primer (2%), nuclease free water (2.5%) and RNA amplicon (20%) were assembled at 4˚C. After being incubated at 65˚C for 2 min and a 10 min incubation at 41˚C, 1.25μl enzyme Mix (Life Sciences NEC-1-24; 25%) was added to the reaction. The reaction took place at 41˚C for 2 h and was then diluted 1:6 into water before applying to the freeze-dried paper devices.

3.2.6 qRT-PCR test

Primers are designed for both GFP gene and 16s rRNA which used as interval control. Bacteria colonies transformed with cognate and non-cognate LIRA and trigger plasmids are inoculated in 6ml LB in triplicates with corresponding antibiotics. All RNAs in bacteria are extracted with commercial RNA miniprep kit (Zymo Research, R2014) following the standard protocol in the manual. Reverse transcription is performed using the commercial kit (Qiagen, 205311) with the protocol from the manual. PCR is taken place with commercial kit (Life Technologies, 4367659) and measured by the Mx3005P qPCR system.

3.3 Results and discussion

3.3.1 LIRA design and validation in vivo

A LIRA is a RNA molecule with a hairpin structure comprised of a large loop and a long stem (Fig. 3.1A). Gene expression from the switch RNA is turned off initially because the ribosome binding site and initial codons of the output gene are enclosed within the duplex of a stem-loop structure. Importantly, the size of the loop in this stem-loop structure is substantially

increased to 21-nts, which is much longer than the conventional loop-mediated riboregulator designs. This large loop provides a large initial docking site for the cognate trigger RNA, facilitating the binding of the trigger RNA and enabling favorable reaction kinetics, which can lead to high dynamic range in ON/OFF ratios. When a trigger RNA is present in the cell, it binds to the exposed bases in the switch RNA loop and proceeds to hybridize with bases within the stem. As this hybridization occurs, the newly formed RNA duplex between the trigger and switch RNA acts as a molecular crowbar that causes the remaining base pairs in the stem to break apart. This stem disruption exposes the ribosome-binding site and initial codons of the output gene, enabling translation of the downstream gene.

The stem length of the LIRA is 27 nucleotides (nt) without considering the clamp domain, which is the base paired domain in the bottom stem below the start codon. This domain is named ‘clamp’ because it can help to lock and seal the start codon within the hairpin structure to prevent signal leakage, when no cognate trigger RNA is present. To insure LIRAs provide low signal leakage at ‘OFF’ state, we designed a series of LIRAs with 0, 3, 6 and 9 base pair (bp) clamps. Testing with flow cytometry, although at least one device of each design exhibits high ON/OFF ratios, the devices from 6 bp designs provided the highest ON/OFF ratios in general (Fig. 3.1B). We thus set 33 nt as stem length and designed 24 different LIRAs via NUPACK.

Those 24 LIRAs were transformed into E. coli BL21 star DE3 together with cognate or non-cognate trigger plasmids and validated by flow cytometry. 16 out of 24 devices exhibit ON/OFF levels over 50-fold, in which 8 of them are over 100-fold, with the highest one reach up to ~350-fold (Fig. 3.1C). The fraction of LIRA designs with ON/OFF > 100 reaches up to around 33%, much higher than the first-generation toehold switches, in which only ~12% devices provided ON/OFF ratios over 100-fold. We expect that further improvements in ON/OFF levels can be obtained from refinements to the LIRA design parameters, for instance through increased loop domain size to increase the thermodynamic driving force and reaction kinetics, and more stringent screening of the secondary structure of the activated trigger/switch complex.

Besides dynamic range, crosstalk level is another considerable parameter to evaluate design strategies, since low crosstalk devices are more likely to be effective for constructing

complex genetic circuits. To test the orthogonality of LIRAs, we selected 16 designs having relatively high ON/OFF ratios and tested all 256 pairwise combinations of switch and trigger RNAs. We used flow cytometry to quantify GFP output from all switch-trigger combinations in triplicate measurements (Fig. 3.1D). Only cognate pairs of LIRAs and RNA triggers provided the highest signal output compared with all the other non-cognate combinations.

Since there are no sequence constraints for trigger RNA sequences, LIRAs thus can be designed to detect arbitrary mRNAs in principle. We designed several mRNA sensors based on one of the high-performance devices. With cognate mRNA triggers, strong GFP signal output was detected in most cases (Fig. 3.1E). This demonstrated that LIRAs can be designed into mRNA sensors to detect mRNAs with completely arbitrary sequences in vivo.

To demonstrate LIRA is compatible with promoters other than T7 promoter and can function well in other bacteria strains, instead of only in E. coli star DE3. We further validated LIRA regulated by different promoters in several other bacteria strains. We tested LIRA #20 with four different promoters in four different bacteria strains, most of them exhibit high dynamic range and low-leakage in flow cytometry tests (Fig. 3.1F). This demonstrates that LIRA can be engineered to function in different cellular environments with excellent performance.

Compared with toehold switches, LIRAs exhibit several advantages. First, since the trigger RNA does not interact with the bottom of the switch RNA stem, LIRAs completely decouple the sequence of the trigger RNA from the bases encoding the output protein. In contrast, toehold switches add at least three undesirable residues to the output gene between the start codon and the linker sequence. These residues can affect output protein folding and stability. Second, the LIRA interaction mechanism results in a system in which a similar number of base pairs exist before and after formation of the trigger/switch complex. This balance in base pairing between these two states results in very sensitive thermodynamics that can be exploited for improved device to device orthogonality[150]. Toehold switches, on the other hand, always gain multiple base pairs upon hybridization to the toehold domain, and thus exhibit less sensitive thermodynamics.

3.3.2 Extremely low leakage in vivo

During the experimental validation, we observed that the 'OFF' state of most LIRAs provided extremely low leakage compared with the 'OFF' state of several toehold switches, and is almost the same as cell autofluorescence (Fig. 3.2A). Rho-independent transcriptional terminators in prokaryotes feature long stem-loop structures that interfere with the progress of the RNA polymerase to halt transcription. While toehold switches have relatively weak stems 15-bp long with a 3-nt bulge, LIRAs have far stronger stems up to 33-nts in length. Given the strength of this stem, we hypothesized that transcriptional regulation could be playing a role in the ultralow leakage we observed in the LIRAs. Thus, we performed qRT-PCR studies on both LIRA #1 and toehold switch N56 to determine the concentrations of the switch RNAs in vivo. We used 16s rRNA as the internal marker for the measurements. At first, we induced the expression of both cognate and non-cognate pairs of switch and trigger for LIRA #1 and toehold switch N56 using three colonies for each. After induction for 3 hours, we performed RNA minipreps to extract cellular RNAs and used reverse transcription to generate cDNA.

Based on the signal curves of SYBR Green from LIRA #1, we can see that expression levels of the housekeeping rRNA are stable for both cognate and non-cognate pairs, but GFP mRNA expression levels are significantly different (Fig. 3.2B). For the cognate RNAs, CT value of GFP cDNA is 18, but it is 20 for non-cognate pair (Fig. 3.2C). For toehold switch N56, CT values for both cognate and non-cognate pairs are 18. The relative gene expressions of different qPCR templates are shown in Figure 3.2D. According to Figure 3.2D, toehold switch N56 does not exhibit regulation at the transcriptional level since the cognate and non-cognate pairs show the same expression levels of GFP transcripts. However, for LIRA #1, cells expressing the cognate trigger/switch pair have 4 times more GFP transcripts than those expressing non-cognate RNA. The relative gene expression levels indicate that LIRAs employ regulation at the transcriptional level. As shown in Figure 3.2D, the CT values of GFP transcripts for the cognate pair of LIRA #1 and toehold switch N56 are the same, indicating no regulation in transcriptional level occurs when cognate pair of switch and trigger appear.

We propose that the trigger RNAs can bind to the newly transcribed switch RNA before the formation of the stem-loop structure. This binding will facilitate the transcription of the whole switch RNA including the downstream GFP gene. However, for non-cognate switch and trigger RNAs, the trigger RNA cannot bind to the newly transcribed switch RNA and the RNA polymerase will be displaced from the DNA due to the formation of the stable stem-loop structure. This finding demonstrates that LIRAs has both transcriptional and translational level regulation.

3.3.3 Paper-based diagnostic with LIRA

The recent outbreak of Zika prompts the need of low-cost and portable diagnostic platforms, which do not require expertise to manipulate and expensive instruments to implement for people living in rural areas. The traditional ways to detect Zika virus are antibody detection, qRT-PCR and isothermal nucleic acid amplification. Antibody detection usually outputs readout very fast, but it provides limited sensitivity. Furthermore, the cross-reactivity from other flaviviruses may result in false positive readouts[89]. qRT-PCR and isothermal nucleic acid amplification provide with more accurate diagnostic readouts since they are targeting at specific regions of virus genomes. However, these techniques require expensive equipment, reagents and experienced expertise to run the test, which may not available for people living in low-resource locations.

Paper-based cell-free systems are a portable and low-cost diagnostic platform for detecting viruses with high sensitivity (Fig. 3.3A). They do not require thermal cycling equipment for running the reaction, and the color change signal can be directly discerned by eye, which does not require additional equipment for detecting the readout. For constructing the paper-based diagnostic platform with LIRA, bacterial cell-free systems together with plasmid encoding for pathogen detecting LIRAs are freeze-dried on a small piece of filter paper disc, which is 2 mm in diameter. When the freeze-dried paper disc is rehydrated with corresponding trigger RNAs, transcription of the synthetic gene network and expression of the regulated reporter protein will be activated to output a color change signal for detection. To improve the performance of LIRA, we

designed a 5' hairpin reconfiguration domain to stabilize the 'ON' state structure, after the hairpin structure is disrupted by the cognate trigger RNAs (Fig. 3.3B).

The pathogen detecting riboregulators displayed high specificity toward their synthetic virus RNA targets with a final concentration of 5 μM (Fig. 3.3C). However, viruses in clinical urine or serum samples are not sufficient to turn on paper-based reactions. To improve the detection sensitivity, an isothermal RNA amplification technique termed NASBA (nucleic acid sequence-based amplification) was incorporated[90] (Fig. 3.3D). In NASBA, reverse transcription takes place with the binding of target RNA template and reverse primer to generate a DNA/RNA duplex. Then the RNase H degrades the RNA temple to release the ssDNA. A forward primer containing T7 promoter initiates the formation of dsDNA. Then the RNA polymerase generates numerous copies of new target RNA templates. Those new RNA templates can react with LIRA and can also serve as new templates for further amplification. To release the RNA template from virus capsid, a simple heating step is needed. After diluted with water and heated up to 95˚C for 2min, sufficient RNA templates can be released for running NASBA.

With NASBA, synthetic RNA templates of Norovirus, YFV with concentrations of 200aM and clinical DENV sample are all successfully detected (Fig. 3.3E). With respect to the clinical DENV sample, we diluted the human serum sample by 10 fold with water and heated it up to 95˚C for 2min before transferring it to NASBA reaction. The total process only takes 3 hours, 2 hours for NASBA reaction and 1 hour for color change reactions. To find out the detection limit of riboregulator based virus sensors, we tested DENV sensor with NASBA products from serial diluted synthetic DENV RNA target ranges from 200 fM to 0.2 aM. And this platform realized the detection of DENV RNA down to 20aM, which equals to 12copies/μl (Fig. 3.3F).

3.3.4 Stem Variants of LIRA

Toehold switches add at least three amino acids between the start codon and the linker region of the regulated gene, which could have deleterious effects on the final output protein. In contrast, no linker is needed and no additional amino acids are required for LIRAs. However, there are six nucleotides located after the start codon in LIRA to help maintain the system at OFF

state when no cognate trigger RNA appears. We term this 6-nt a 'clamp' as it is designed to clamp down on leakage from the devices. We hypothesized that since the clamp does not interact directly with the trigger RNA, it should be possible to use the first 6 nucleotides of any gene as a successful clamp. To test this hypothesis, we designed 4 new variants from LIRA #1. These four switches have the same sequence as #1, but the clamp sequences are randomly designed. Then we transformed each switch with cognate trigger #1 and each switch with one non-cognate trigger into E. coli and measured the GFP output after inducing by IPTG. The ON/OFF ratios from the mode GFP fluorescence value measured by flow cytometry are shown in Figure 3.4A.

As Figure 3.4A shows, although variance exists in the output signals of each switch-trigger combination, all systems are effectively turned on with ON/OFF > 50. Based on Figure 3.4B and 3.4C, output variance are resulted from ON states, since all OFF states are almost the same. The high ON/OFF ratios indicate that clamp has no sequence constraints and thus reasonable performance can be expected for LIRAs regulating diverse output genes. We also designed four switches with different clamps from LIRA #5, and they too exhibit similar performance. Therefore, with no sequence constraints of the clamp, we can use the first 6 nucleotides of any gene as clamp. We attribute the observed variation in output from the LIRAs to changes in the secondary structure of the trigger/switch activated complex, which can affect translational efficiency, and the effect of the changes in clamp sequence to the folding of the output GFP.

3.3.5 Loop-initiated RNA repressors

All the above riboregulators are designed for turning on expression of regulated genes by binding with the trigger RNA. However, to precisely control gene expression and make more complex genetic circuits, another type of riboregulator termed 'repressor' has been studied by many researches[151]. For repressors, gene expression is initially in its 'ON' state. As is shown in Figure 3.5A-B, RBS domain and AUG start codon are located at a linear structure in upstream of regulated gene, so the regulated GFP gene can be translated initially. In the hairpin structure, we rationally designed a domain that is completely complementary to RBS domain. To weaken the

transcriptional-level regulation observed in LIRAs, we designed a structure containing two small hairpins, each of the hairpins has a relatively shorter stem. For the designs shown in Figure 3.5A and 3.5B, the first hairpin is 21 nts in length and the second one is 14 nts in length. The design in Figure 3.5B has a 12 nts 5' toehold domain, which is designed for enhancing the transcription of the whole switch RNA. The data of the best device from each design is shown in Figure 3.5C, which can achieve over 12-fold repression.

3.4 Conclusion

LIRAs are a new type of translational regulator that exploit an effective loop-mediated interaction with the trigger RNA. These de-novo-designed riboregulators provide high dynamic range, low crosstalk and extremely low leakage in vivo. Since no correlated domains with predefined sequences located in the trigger RNA binding domain, LIRAs can be engineered to detect RNAs with completely arbitrary sequences. By incorporating with paper-based diagnostic platform, LIRA realized the detection of viruses down to 20aM, which is equals to 12 copies/μl.

Figures

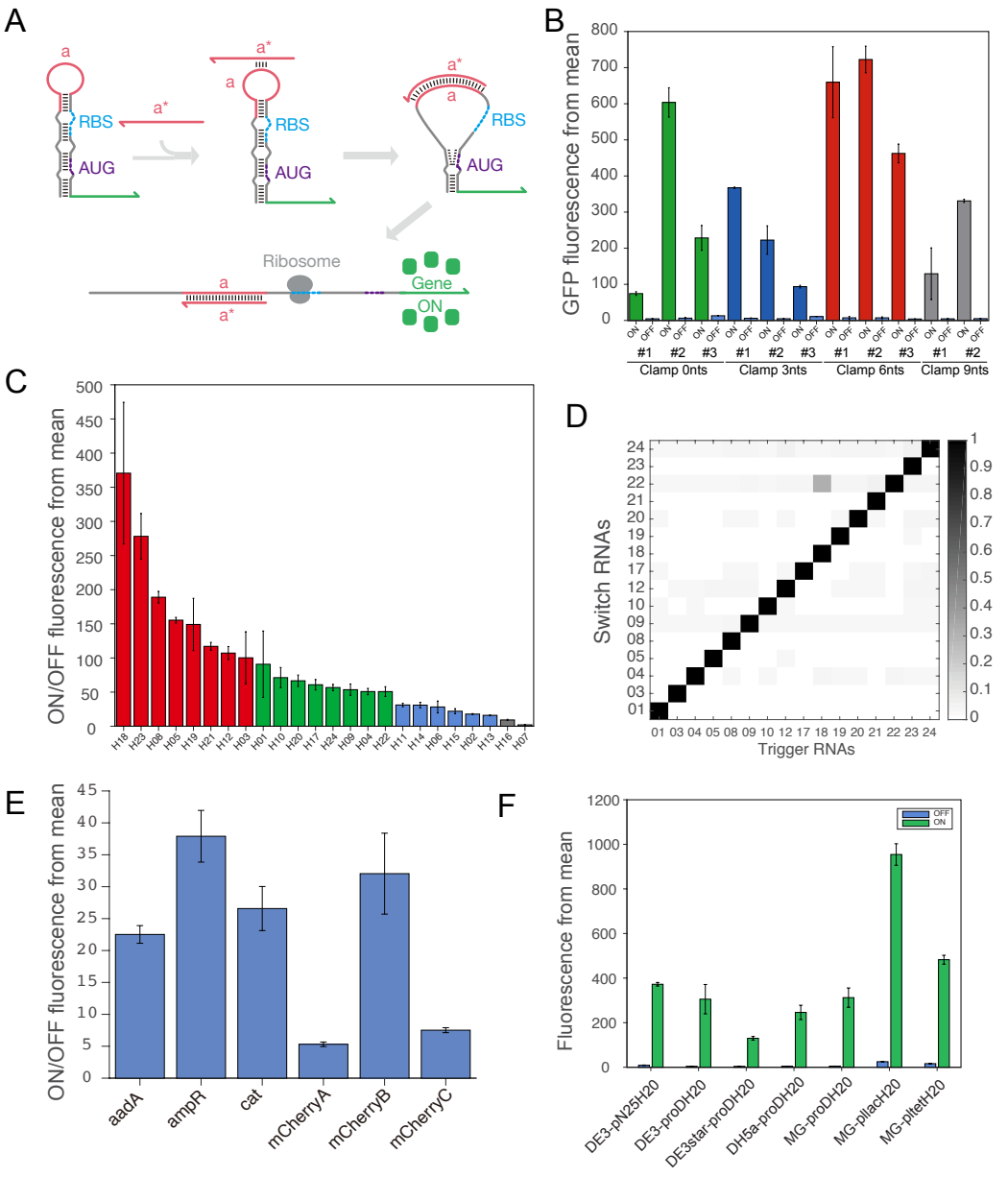
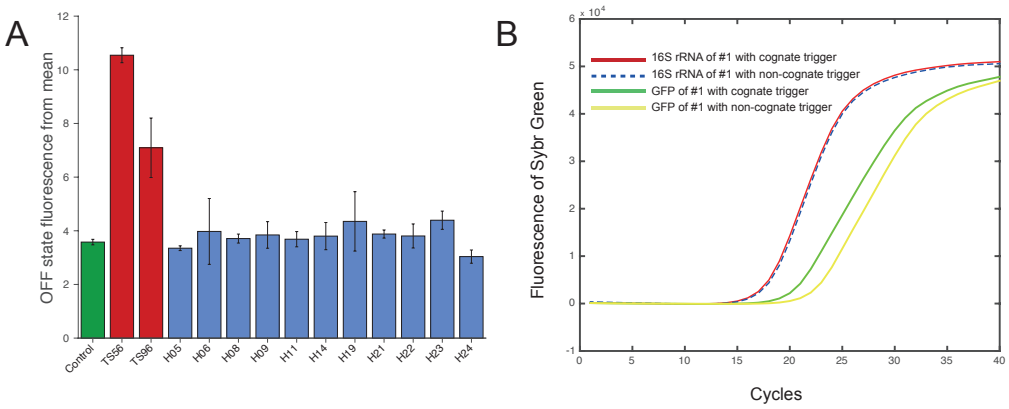


Figure 3.1 Structure and in vivo validation of LIRAs.

(A) Structure and interaction mechanism of LIRA; (B) Flow cytometry measurement of LIRA designs with different clamp length; (C) ON/OFF ratios of 24 different riboregulators; (D) Orthogonal test for 16 pairs of riboregulators and trigger RNAs; (E) In vivo test of LIRA based mRNA sensors; (F) LIRA tests in different bacterial strains with different promoters.



C

Average CT value	1# Cognate	1# Non-cognate	Toehold 56# Cognate	Toehold 56# Non-cognate
16S rRNA	17	17	17	17
GFP	18	20	18	18

D

	ΔCT	-ΔΔCT	relative gene expression
#1 Cognate	1	0	1
#1 Noncognate	3	-2	0.25
Toehold N56 Cognate	1	0	1
Toehold N56 Noncognate	1	0	1
#1 Cognate	1	0	1
Toehold N56 Cognate	1	0	1

Figure 3.2 Extremely low leakage of LIRAs in vivo.

(A) Leakage comparison of toehold switches and LIRAs; (B) qRT-PCR test of LIRA #1; (C) Cycle threshold (CT) values of 16S rRNA and GFP mRNA in different conditions; (D) Delta CT value of each codition.

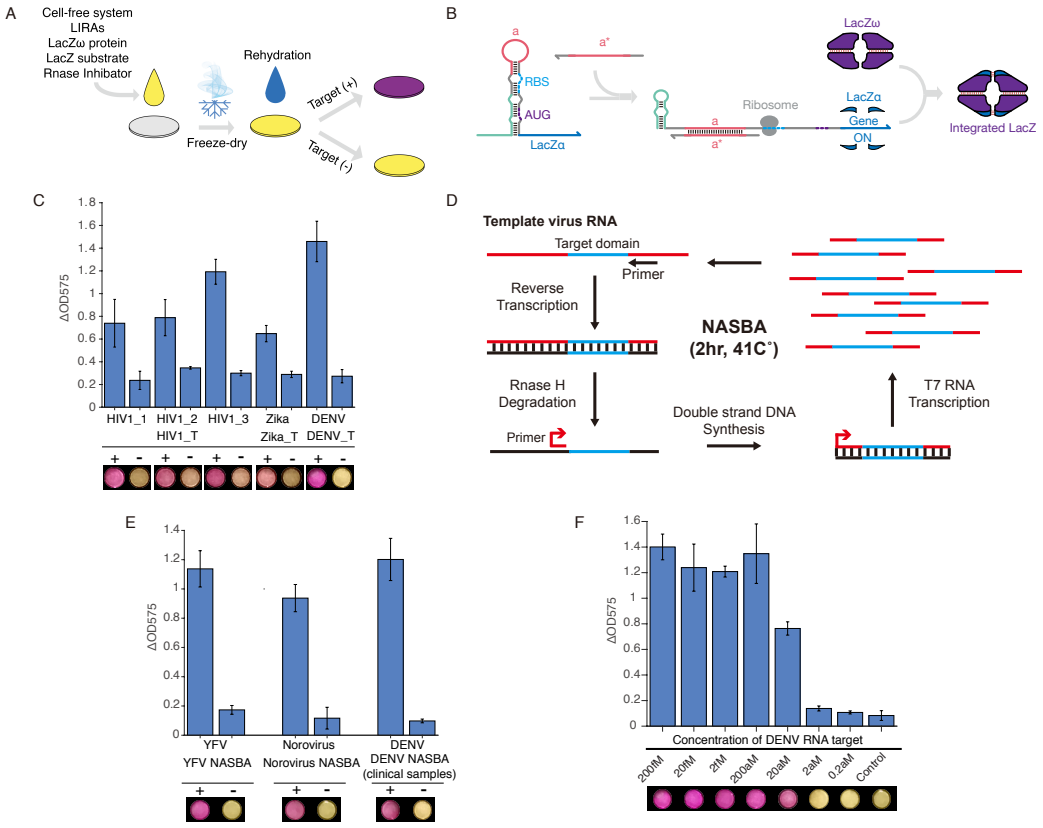


Figure 3.3 Pathogen-detecting LIRAs with paper-based diagnostic platform.

(A) Schematic of paper-based diagnostic platform; (B) Design of pathogen detecting LIRAs; (C) Synthetic RNA detection, photos were taken after 80mins; (D) Schematic of NASBA process; (E) NABSA product detection, photos were taken after 1.5 hours for YFV and Norovirus detection, 2 hours for DENV clinical sample detection; (F) Detection limit test of DENV sensor, photos were taken after 1.5 hours.

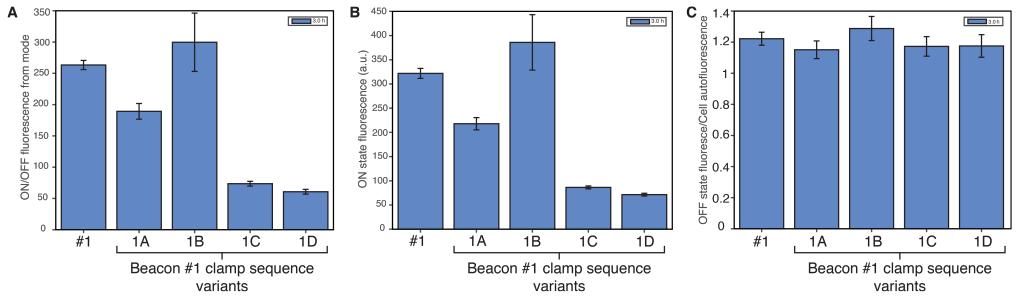


Figure 3.4 LIRA clamp variant test.

(A) ON/OFF ratios for the four switches designed from LIRA #1. All four switches have the same cognate trigger; ON state were measured from cognate pairs of switch and trigger; OFF states were measured from non-cognate pairs of switch and trigger; (B) ON state/Cell autofluorescence ratios for the four switches designed from LIRA #1. ON state were measured from cognate pairs of switch and trigger; (C) OFF state/Cell autofluorescence ratios for the four switches designed from LIRA #1. OFF states were measured from non-cognate pairs of switch and trigger.

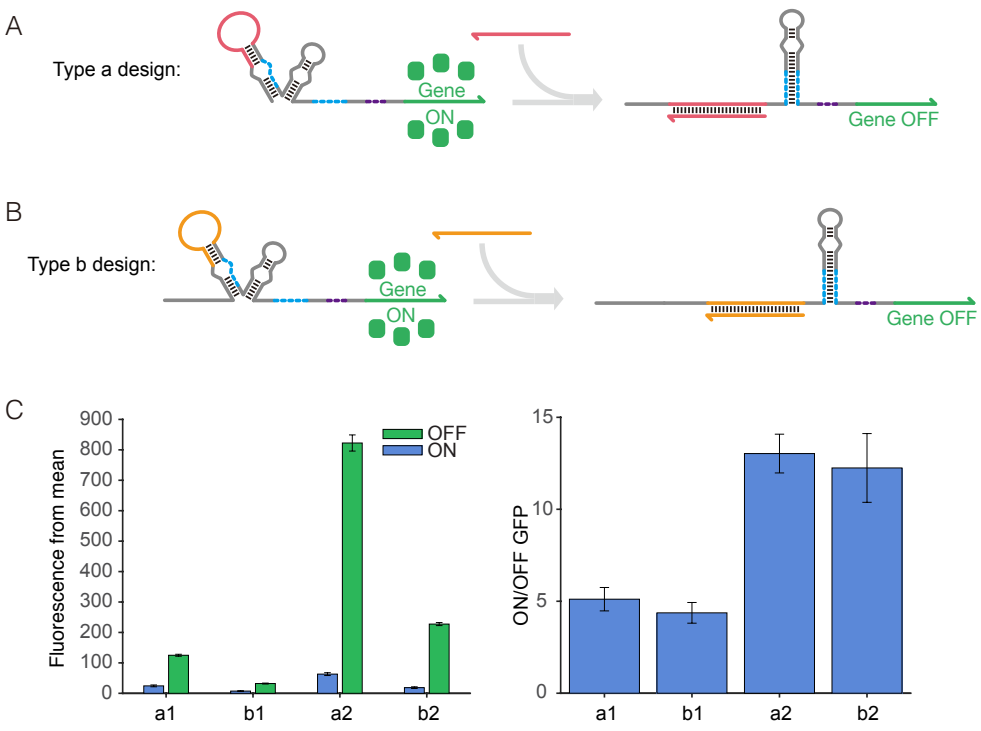


Figure 3.5 Different repressor designs and in vivo validation.

(A-B) Two different types of rationally designed beacon repressors; (C) The data of best devices from each of those two type designs, ON state and OFF state respectively (left) and ON/OFF ratios (right).

CHAPTER 4

LIRA-BASED RIBOCOMPUTING DESIGNS

4.1 Introduction

Researchers have developed numerous genetic circuits with logic computing functions aiming at diverse applications[16,18,21,23,24,32,56,152,153]. Those genetic circuits were designed with creative ideas and took advantages of versatile interactions among different biological parts, which can fulfill their design purposes effectively. However, those logic circuits were usually designed with multiple layers, and need to incorporate numerous proteins or enzymes, which requires extensive resources for carrying out logic computing functions. Moreover, the limited number of orthogonal biological parts is still a challenge for constructing more complicated genetic circuits. RNA-only nanodevice is a promising design to overcome those limitations because of their high programmability. With the predictable base-pairing rules, complex structures with logic computing functions could be designed *in silico* with versatile software and online servers, such as NUPACK. As a single-layer device, input transduction, computation and signal output are all implemented in a self-assembled RNA molecule, which improves the processing efficiency of genetic circuits.

Recently, toehold switch based RNA-only nanodevices with diverse logic computing functions were developed[85]. The de-novo-designed single layer, RNA-only nanodevice, can be designed to carry out multiple sophisticated logic computations in vivo. By arranging six different toehold switches in a row, ‘Or’ logic nanodevices with up to six different inputs are successfully constructed. However, the output protein of each input is different, since the RBS domains located in upstream hairpins can initiate the translation of downstream toehold switches into peptides attached to the reporter protein, which may interfere with proper protein function. Moreover, the arrangement of several toehold switches in a row leads to higher protein expression levels for downstream hairpins compared with the upstream ones, since the strong hairpin secondary structures can impede the translating ribosome. For designing ‘And’ logic nanodevices, the trigger RNA is evenly split into several parts with complementary domains added onto each of them. When being transcribed together, those split triggers can cooperatively

interact with each other to form the full-length trigger, which then binds with toehold switch to turn on gene expression. With this mechanism, ‘And’ logic nanodevice with up to four different inputs were successfully tested, but those split RNA inputs are not independent, which hinders its applications for RNA detection and further genetic circuit constructions.

To overcome those limitations, we developed LIRA-based ribocomputing nanodevices with logic-computing functions. By rationally fusing multiple LIRAs into one nanodevice and arranging the positions of RBS and start codon, ribocomputing nanodevices with ‘Or’ and ‘And’ logic operations were successfully constructed and validated in Eschericia coli and paper-based cell-free systems.

4.2 Methods and materials

4.2.1 Strains and growth conditions

The following E. coli strains were used in this study: BL21 Star DE3 (F− ompT hsdSB (rB−mB−) gal dcm rne131 (DE3); Invitrogen), BL21 DE3 (F− ompT hsdSB (rB−mB−) gal dcm (DE3); Invitrogen), MG1655Pro (F− λ− ilvG- rfb-50 rph-1 SpR lacR tetR), and DH5α (endA1 recA1 gyrA96 thi-1 glnV44 relA1 hsdR17(rK−mK+) λ−; Invitrogen). All strains were grown in LB medium at 37 °C with appropriate antibiotics.

4.2.2 Plasmid construction

Plasmids were constructed by PCR and Gibson assembly. Single-stranded DNAs for expressing logic gate and target RNAs were purchased from Integrated DNA Technologies and amplified into double-stranded DNA via PCR. The amplified DNAs were then connected with plasmid backbones by 30-bp homology domains using Gibson assembly. All Gibson assembly products were transformed in the E. coli DH5α strain and then sent out for sequence validation via Sanger sequencing. Backbones used for constructing the plasmids were amplified from the commercial vectors pET15b, pCOLADuet, and pCDFDuet (EMD Millipore) via PCR followed with Dpn1 treatment. The reporter protein for all plasmids is GFPmut3b with an ASV degradation tag.

4.2.3 Flow Cytometry Measurements and Analysis

Bacteria colonies transformed with logic gate and target RNA plasmids are inoculated in 1ml LB in triplicates with corresponding antibiotics. On the second day, 5ul overnight cultured medium is diluted by 100 folds in 495ul fresh LB with 60% kanamycin, 50% ampicillin and 50% spectinomycin. After 80mins recover, IPTG is added into each well with a final concentration of 0.1mM. And flow cytometry measurement is performed after 3, 4 and 5 hours inducement.

Flow cytometry was performed using a S1000 cell analyzer (Stratedigm) equipped with a high-throughput auto sampler (A600, Stratedigm). Before running measurement, cells were diluted by ~10 folds into phosphate buffered saline (PBS) in 384-well plates. Forward scatter (FSC) was used for trigger, and 40,000 individual cells were recorded. Cell populations were gated according to their FSC and side scatter (SSC) distributions as described previously. The GFP fluorescence signal outputs of these gated cells were used for following calculations. Error levels for the fluorescence measurements of ON state and OFF state cells were calculated from the SD of measurements from at least three biological replicates. The relative error levels for the ON/OFF fluorescence ratios were then determined by adding the relative errors of ON and OFF state fluorescence in quadrature.

4.2.4 Cell-Free Reactions

Cell-free transcription-translation systems (NEB, PURExpress) were prepared for freeze-drying according to following recipe: cell-free solution A, 40%; cell-free solution B, 30%; RNase Inhibitor (Roche, 03335402001, distributed by MilliporeSigma), 2%; chlorophenol red-b-D-galactopyranoside (Roche, 10884308001, distributed by MilliporeSigma, 24mg/ml), 2.5%; with the remaining volume reserved for LIRA-based ribocomputing switch DNA, water and lacZα peptide added to a final concentration of 2 μM. When testing ribocomputing devices with paper-based system, the final concentrations of the plasmids are 15ng/μl.

Filter paper (Whatman, 1442-042) for deposit and freeze-drying the cell-free system was first blocked with 5% bovine serum albumin (BSA) overnight. The paper was washed three times in water for 5 to 10 min after overnight blocking. The paper was transferred on a hot plate at 50˚C

for drying and then cut into 2mm diameter paper disks with a biopsy punch. The disks were then transferred into 200μl PCR strips and 1.8μl of the above cell-free reaction mix was applied to each of them. Liquid nitrogen was used for freeze the PCR strips containing those paper devices. The frozen paper disks were dried overnight with a lyophilizer. Plate reader tests were carried out on the freeze-dried paper disks 2-4 days later. With the systems stored in nitrogen environment, shielded from light and together with the silica gel desiccation packages as described previously, the paper disks remained active for at least a month under room temperature.

4.2.5 NASBA

NASBA experiments were carried out following the standard protocols: reaction buffer (Life Sciences, NECB-24; 33.5%), nucleotide mix (Life Sciences NECN-24; 16.5%), RNase inhibitor (Roche, 03335402001; 0.5%), 12.5 μM of each DNA primer (2%), nuclease free water (2.5%) and RNA amplicon (20%) were assembled at 4˚C. After being incubated at 65˚C for 2 min and a 10 min incubation at 41˚C, 1.25μl enzyme Mix (Life Sciences NEC-1-24; 25%) was added to the reaction. The reaction took place at 41˚C for 2 h and was then diluted 1:6 into water before applying to the freeze-dried paper devices.

4.3 Results and discussion

4.3.1 LIRA-based ‘Or’ logic gates

A 2-input ‘Or’ logic nanodevice comprises two different LIRAs, and the RBS and start codon are localized in the shared bottom stem, so each RNA input can disrupt the bottom stem to turn on gene expression (Fig. 4.1A). A hairpin reconfiguration domain was inserted between those two LIRA hairpins to maintain a stable ‘ON’ state when the upstream hairpin is disrupted. And strong GFP expressions at similar levels were detected for all logical TRUE conditions tested via flow cytometer (Fig. 4.1B).

A 3-input ‘Or’ logic gate nanodevice is constructed in the same way with two hairpin reconfiguration domains inserted to each of the upstream hairpins (Fig. 4.1C). Each single input can disrupt its corresponding hairpin under the help of the small hairpin reconfiguration domain to

release the blocked RBS and start codon from the bottom stem. The output signals were detected for all logical TRUE conditions, but the signals detected from single-input cases are weaker than those with 2 to 3 inputs (Fig. 4.1D). This may be resulted from the non-fully disruption of the structure, as the entire tightly sealed structure is too stable for single RNA input to completely disrupt. When two or more RNA inputs work together, the entire structure can be totally unwound.

4.3.2 LIRA-based 'And' logic gates

To design nanodevice with 2-input 'And' logic computation function, besides fusing two different LIRAs together, the RBS and start codon are also arranged into the last LIRA hairpin structure (Fig. 4.2A). In this case, the disruption of upstream LIRA by the RNA input cannot release the RBS and start codon. But without the disruption of the upstream LIRA, the RNA input for the last LIRA cannot disrupt the entire tightly sealed nanostructure. Only when the two RNA inputs are transcribed together, the expression of reporter protein can be turned on. As demonstrated by the flow cytometry measurement, strong GFP expression was detected with the logical TRUE condition, and no signal outputs were detected for all logical FALSE conditions (Fig. 4.2B). This 2-input 'And' logic design does not require a hairpin reconfiguration domain. We also tested a 2-input 'And' logic nanodevice with a hairpin reconfiguration domain inserted, but it shows lower ON/OFF level compared with this one.

A 3-input 'And' logic nanodevice is also constructed and inserted with two independent hairpin reconfiguration domains (Fig. 4.2C). Each RNA input alone or any pair of two RNA inputs cannot disrupt the final LIRA to release the blocked RBS and start codon. When all three RNA inputs work together to disrupt each hairpin sequentially, strong GFP signal output can be detected. And from the flow cytometry measurement, we can see that the correct truth table was successfully achieved (Fig. 4.2D).

4.3.3 LIRA- and toehold-switch-based 'And' logic gates

Although LIRA-based ribocomputing devices solved the limitations of toehold-switch-based ones, the design difficulty for LIRA-based devices is much harder than toehold switch ones.

LIRA-based logic gates usually requires more complex secondary structures with the increasing numbers of inputs, which will result in stringent sequence requirements during designing. Since it is very useful to get a design strategy that can combine the advantages from both toehold switch and LIRA design strategies, we further validated a new type of 'And' logic gates, which we named LIRA-Toehold 'And' logic gates.

We first validated a 2-input 'And' logic gate device (Fig. 4.3A). In this design, no hairpin reconfiguration domain is required, and toehold domain from toehold switch is blocked in the stem of LIRA, just at the position where RBS was located previously. Trigger for LIRA can disrupt the hairpin structure to release the toehold domain accessible for toehold trigger binding. However, toehold trigger itself cannot unwind the entire structure, since the toehold domain is tightly sealed in the LIRA stem, which is demonstrated by flow cytometry measurement (Fig. 4.3B).

Then we constructed a 3-input 'And' logic device (Fig. 4.3C). In this design, the 5' toehold switch has a completely changed shape compared with normal toehold switches. Its loop domain is completely removed, and only leaves the binding domain together with toehold domain. Its stem is just a single-stranded DNA, which partially base-paired with part of repressed gene sequences. Testing with flow cytometry, we got a correct truth table (Fig. 4.3D).

4.3.4 Paper-based diagnostic with logic gates

For previously developed paper-based diagnostic platforms, each paper device can only detect one type of virus, which hinders its application in detecting multiple viruses or different virus subtypes. Moreover, each sensor only targets a small region in the virus genome, which may not be reliable for diagnostics that require stringent accuracies. We thus investigated whether the logical computation ability of LIRA-based ribocomputing nanodevices can be incorporated into diagnostic applications. In principle, 'Or' logic nanodevice should be able to carry out the detection of multiple viruses or different subtypes of one virus on the same paper device, and 'And' logic nanodevice should provide more accurate readouts by targeting at multiple domains in the virus genomes.

To demonstrate this, different subtypes of HIV1 were selected for ribocomputing nanodevice designs, since HIV is regarded as one of the major worldwide health concerns and HIV1 is the predominant virus among all HIV infections worldwide[154]. There are four groups belong to HIV1, group M, N, O and P, and group M is the major one that responsible for the global HIV epidemic[155]. Under group M, there are 9 different subtypes, subtypes A, B, C, D, F, G, H, J and K. We chose subtypes A, B and C for logic ribocomputing studies, since they account for over 70% infections of HIV1 worldwide[155].

We downloaded and aligned 114 sequences for subtype A, 81 sequences for subtype B and 200 sequences for subtype C from NCBI database and got conserved sequences for each subtype, respectively. We also found a region that conserved in all those 3 subtypes. Based on the conserved sequences, we designed a 2-input 'Or' logic nanodevice targeting at HIV1 subtypes B and C (Fig. 4.4A) and a 2-input 'And' logic nanodevice targeting at the conserved domain in subtype B and the conserved domain found in all three subtypes (Fig. 4.4C).

Unlike the in vivo nanodevices, we made some modifications to the nanodevice structures. With respect to 'Or' logic nanodevice, we set the binding region covers the full loop and stem (Fig. 4.4A). For 'And' logic nanodevice, we inserted a hairpin reconfiguration domain to the structure and designed a smaller loop for the downstream hairpin (Fig. 4.4C). RNA-keys are amplified by NASBA with synthetic RNA as templates, and applied to paper devices freeze-dried with either 'Or' or 'And' ribocomputing nanodevices. For 2-input 'Or' logic gate, NASBA products from subtype B or subtype C can turn on the system efficiently demonstrated by the color change (Fig. 4.4B). For 2-input 'And' logic gate, color change can only occur when both NASBA products from subtype B are being applied on to the paper device (Fig. 4.4D).

To better illustrate our paper-based diagnostic platform with logic computing functions, we further developed sensors targeting at hemagglutinin (HA) and neuraminidase (NA) genes of different influenza A subtypes. The influenza A subtypes that contain same HA or NA genes show high similarity in RNA sequences, but show completely different sequences with irrelevant HA or NA genes (Fig. 4.5A). Since conventional single input riboregulator is not able to distinguish one from another, the two-input 'and' logic strategy can be incorporated by targeting at both HA and

NA genes at the same time when implement influenza A viruses genotyping. We developed 3 sensors targeting at H5N1, H1N1 and H1N2 subtypes, and tested them with either single cognate RNA input or both cognate RNA inputs. Only when both cognate RNAs were added to the paper device, can we detect color change reaction (Fig. 4.5B). Both RNA inputs from different subtypes was not able to induce the paper-based reaction. Thus, with 2-input 'and' logic devices, we can realize more accurate diagnostic or even in subtypes genotyping applications.

4.4 Conclusion

The functional 'Or' and 'And' logic ribocomputing nanodevices provide excellent in vivo performance with low leakage at 'OFF' states from logical FALSE conditions and high 'ON' state signals with logical TRUE conditions. During NUPACK design and experimental screening, we found that the prescribed structure in the design script is more important for getting functional devices, especially for the designs of ribocomputing devices. With improper parameters set for prescribed structures, such as stem and loop lengths, NUPACK package can generate designs that will probably fail in experimental tests. If the prescribed structure is suitable, working devices can always be screened from multiple designs.

Figures

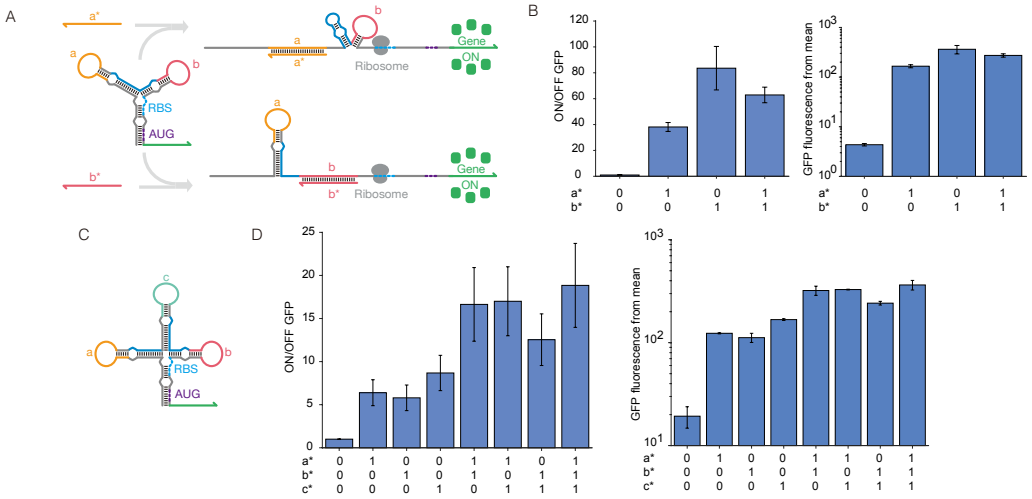


Figure 4.1 Structure and in vivo validation of LIRA-based ‘or’ logic.

(A) Design of 2-input ‘Or’ logic gate and schematic of interaction with triggers; (B) In vivo performance of 2-input ‘Or’ logic gate, ON/OFF ratios (left) and logic scale on state signal (right); (C) Design of 3-input ‘Or’ logic gate; (D) In vivo performance of 3-input ‘Or’ logic gate, ON/OFF ratios (left) and logic scale on state signal (right).

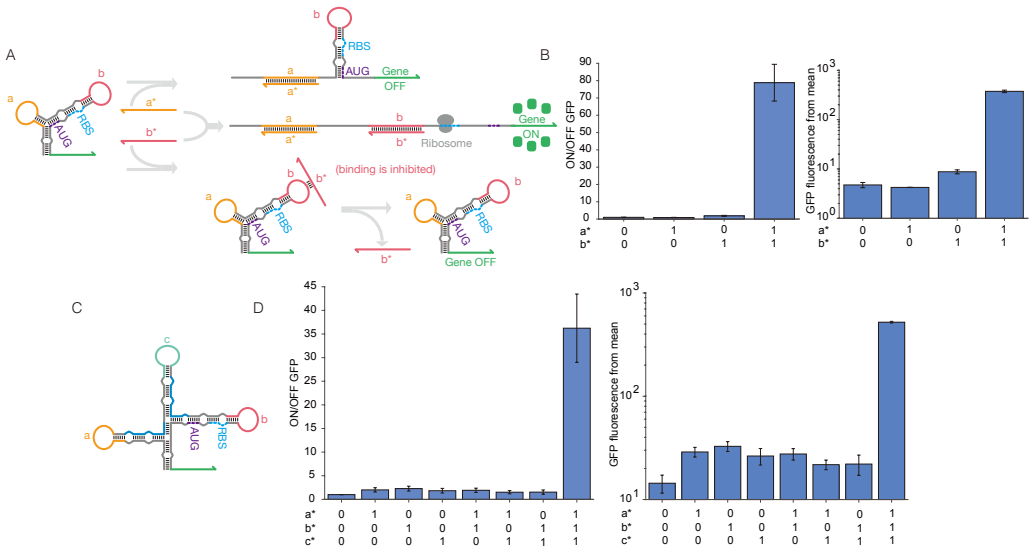


Figure 4.2 Structure and in vivo validation of LIRA-based ‘and’ logic.

(A) Design of 2-input ‘And’ logic gate and schematic of interaction with triggers; (B) In vivo performance of 2-input ‘And’ logic gate, ON/OFF ratios (left) and logic scale on state signal (right); (C) Design of 3-input ‘And’ logic gate; (D) In vivo performance of 3-input ‘And’ logic gate, ON/OFF ratios (left) and logic scale on state signal (right).

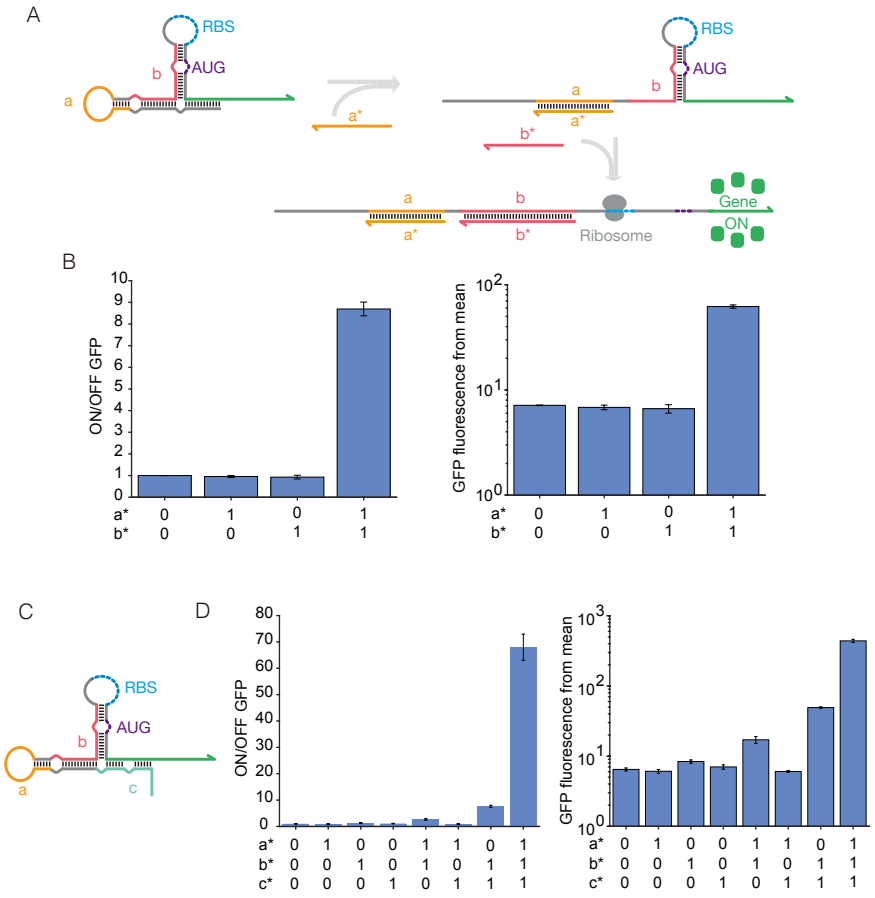


Figure 4.3 Structure and in vivo validation of LIRA- and toehold-switch-based ‘and’ logic. (A) Design of 2-input ‘And’ logic gate and schematic of interaction with triggers; (B) In vivo performance of 2-input ‘And’ logic gate, ON/OFF ratios (left) and logic scale on state signal (right); (C) Design of 3-input ‘And’ logic gate; (D) In vivo performance of 3-input ‘And’ logic gate, ON/OFF ratios (left) and logic scale on state signal (right).

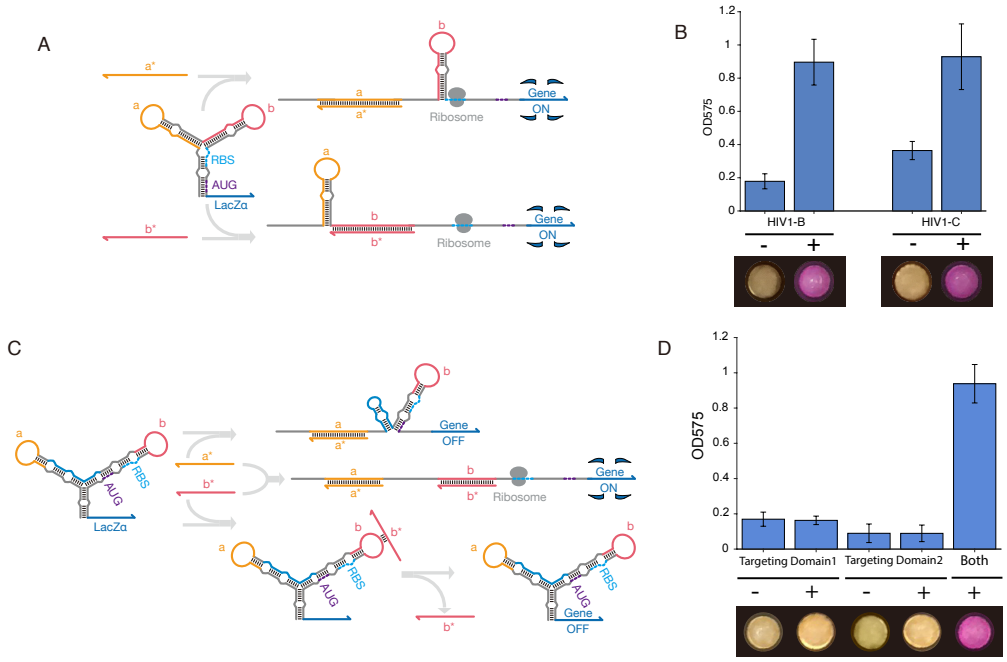


Figure 4.4 Logic gates tests with paper-based diagnostic platform.

(A) Design of 2-input 'Or' logic gate for HIV1-B and HIV1-C detection; (B) Detection of HIV1-B and HIV1-C, photos were taken after 1.5 hours; (C) Design of 2-input 'And' logic gate for HIV1-B differentiation; (D) Differentiation of HIV1-B from HIV1-A and HIV1-C, photos were taken after 2.5 hours.

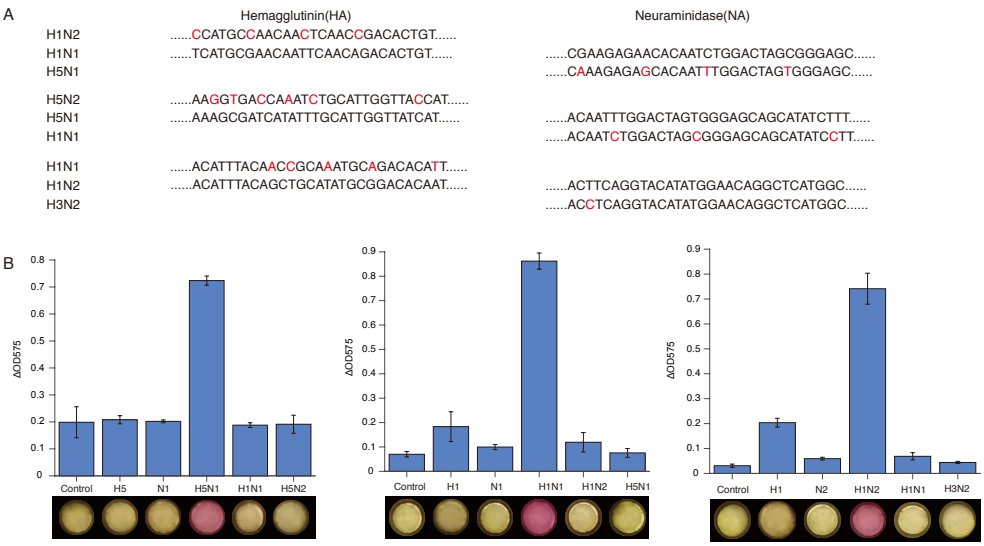


Figure 4.5 Influenza A subtypes genotyping with paper-based diagnostic platform. (A) Targeting sequences of different influenza A subtypes, the mismatched bases are highlighted with red color. (B) Paper-based detection of different influenza A subtypes.

CHAPTER 5

CONCLUSION AND FUTURE DIRECTIONS

The fundamental building blocks for constructing complex synthetic gene networks are effective biological parts with wide dynamic range and low crosstalk. However, the limited number of orthogonal parts still hinders the construction of effective biological networks. RNA-based components are promising solutions to solve this problem, since their predictable base pairing rules, which also provides RNA-based parts with higher programmability than protein-based parts. The toehold switch is a breakthrough in RNA-based component designs. By blocking both RBS and start codon in a hairpin structure but without binding to these sequences, gene expression can be regulated by a short trigger RNA molecule that can adopt nearly any sequence. By rationally modifying the design parameters of toehold switch, such as loop size and toehold length, the ON/OFF ratios of toehold switches can be tuned accordingly, with the ON/OFF ratio of the best toehold switch reaches over 600-fold. Besides its high dynamic range, the toehold switch also provides extremely low crosstalk between different designs. Moreover, toehold switch based complex ribocomputing devices also perform well in vivo, which demonstrated the high programmability of RNA-based components. Since no sequence constraints exist in trigger RNAs, the toehold switch is developed to detect RNA molecules with completely arbitrary sequences. Viruses with RNA genomes, like Ebola and Zika, were all successfully detected by incorporating with paper-based diagnostic platform. To detect clinical samples, in which virus concentration is very low, isothermal amplification reaction, such as NASBA, is incorporated. With target RNA amplification by NASBA, paper-based diagnostic platform successfully detected Zika virus in clinical samples.

With this platform, we turned to see whether we could expand the paper-based platform to detect other viruses. Norovirus can cause human gastroenteritis and lead to over 200,000 deaths in developing countries annually, so a portable and low-cost detecting method is required urgently by people living in rural areas. Compared with traditional detection methods, the paper-based diagnostic platform exhibits several advantages. First, it does not require any expensive instrument, such as thermal cycler and plate reader, to either carry out the reaction or detect the

reaction readout. Second, it is portable and low-cost, since all components are freeze-dried onto a small paper disk, and reactions can be activated by simply rehydration with RNA target. Third, only specifically amplified RNA targets can turn on the color change reaction, since the amplified RNA targets need to interact with the sensor, whose sequence is pre-defined. This mechanism completely prohibits false positive readouts. However, unlike antibody-based detection, paper-based diagnostic platform usually requires longer times, which includes 2-hr isothermal amplification and over 1-hr paper-based reaction. To shorten the time cost, we can either shorten the reaction time for isothermal amplification or decrease the time cost for paper-based reaction. But a 2-hr isothermal amplification is required for amplifying samples containing extremely low virus concentrations. Although virus concentrations can reach up to femtomolar range in some clinical samples, most of them only contain viruses in picomolar or attomolar ranges, which requires 2 hr of isothermal amplification to produce enough RNA targets for paper-based reactions.

To speed up the process, we began to improve the paper-based color change reactions to see whether we could find a proper way to shorten the time cost. Since lacZ is a relatively long reporter protein compared with GFP and mCherry, the time cost should be decreased if we can replace it with shorter ones. Thus, alpha-complementation of lacZ is incorporated into our designs. Each lacZ monomer contains two subunits, lacZα and lacZω, and lacZα is ~50 amino acids long while lacZω is ~970. Without lacZα, lacZω itself cannot form active tetramer. With this mechanism, we replaced lacZ gene with the short DNA oligonucleotides encoding for lacZα peptides, and freeze-dry the lacZω protein together with all the other components that required for paper-based diagnostic system. Since lacZα is much shorter than lacZ, the time requirement during transcription and translation is decreased. And the folding of lacZ protein is also accelerated, because the large subunit lacZω is pre-synthesized and being freeze-dried onto paper disks directly, and the folding of lacZω has been done during synthesis. With this improvement, we got 41% reduction for time cost during paper-based reactions.

Although toehold switch provides excellent performance, it has its own limitations, such as noticeable signal leakage in vivo, unwanted amino acids being translated with reporter protein

and sequence constraints of in-frame stop codon. To overcome the above limitations, we developed a new type of de-novo-designed riboregulator, known as a LIRA, which also shows wide dynamic range and low cross talk for regulation of gene expression. We designed 24 devices, and 33.3% of them showed ON/OFF ratio higher than 100, which is larger than the percentage observed for the first-generation toehold switches. Combined with our paper-based colorimetric diagnostic platform and cell-free system, LIRAs were engineered for detecting a variety of virus RNAs, such as HIV, Zika, DENV, Norovirus and YFV. Among those pathogen detections, the detection limit of DENV can reach down to 20aM, which equals to ~12 copies/μl sample.

Toehold-switch-based ribocomputing devices also have their own limitations. For ‘Or’ logic devices, each input will induce the translation of downstream toehold switches into peptides together with reporter proteins. The secondary structure of downstream toehold switches affect the protein expression induced by upstream inputs. With respect to ‘And’ logic, inputs are not independent since all RNA triggers need to be assembled together to form the full-length trigger first, and then turn on gene expression by interacting with toehold switch. All above limitations hinder their applications in many aspects. To overcome those limitations, we developed LIRA-based ribocomputing devices. For ‘Or’ logic, each input has exactly same output proteins, without adding any unrelated amino acid to output proteins. Protein expression level is at similar level with the same amount of RNA inputs. For ‘And’ logic, each input is completely independent, which makes it possible for detecting multiple RNA molecules with arbitrary sequences at the same time. Based on this, we also constructed LIRA-based ribocomputing devices into virus RNA sensors to expand the use of paper-based diagnostic platform.

RNA-based biological parts exhibit many advantages over protein- and promoter-based biological parts, which usually take complex layered designs and cost substantial cell resources for carrying out their functions. RNA-only devices thus can be engineered for versatile applications in many aspects. For in vivo studies, more complex biological circuits can be constructed via either toehold or loop-mediated RNA-RNA interactions. Toehold switch and LIRA are all translational regulating designs, but they can be easily engineered into transcriptional level

controls by regulating the formation of transcriptional terminators during transcription. By constructing transcriptional level RNA regulators, in which RNA input can regulate the transcription of RNA output, RNA-only and layered circuits can be realized. Since RNA is more programmable than protein based circuits, complex biological circuits with multiple layers will become much easier with RNA based *in silico* designs. Another important point for LIRA is that there is completely no sequence constraint for LIRAs, so it is more compatible with RNA aptamer and endogenous mRNA and miRNA regulations than toehold switches. LIRA can thus be incorporated into either RNA imaging, gene regulation in mammalian cells or even in therapeutic studies. With respect to in vitro application, diagnostic may still be the major point. But those RNA-based designs can be engineered with different outputs instead of using color change output signals. This might be helpful to improve sensitivity of the RNA detecting system and decrease the time cost for detecting the readouts, which is the major drawback for the paper-based diagnostic platform.

In summary, RNA based biological parts are promising building blocks for constructing complex biological circuits with versatile functions. Considering its high programmability and degradable properties, RNA based designs can serve human with a promising ‘key’ to solve complicated ‘locks’ in the future.

REFERENCES

1 Endy, D. Foundations for engineering biology. *Nature* **438**, 449-453, doi:10.1038/nature04342 (2005).

2 Benner, S. A. & Sismour, A. M. Synthetic biology. *Nat Rev Genet* **6**, 533-543, doi:10.1038/nrg1637 (2005).

3 Khalil, A. S. & Collins, J. J. Synthetic biology: applications come of age. *Nat Rev Genet* **11**, 367-379, doi:10.1038/nrg2775 (2010).

4 Monod, J. & Jacob, F. Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb Symp Quant Biol* **26**, 389-401 (1961).

5 Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nat Rev Microbiol* **12**, 381-390, doi:10.1038/nrmicro3239 (2014).

6 Ideker, T. *et al.* Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **292**, 929-934, doi:10.1126/science.292.5518.929 (2001).

7 Westerhoff, H. V. & Palsson, B. O. The evolution of molecular biology into systems biology. *Nat Biotechnol* **22**, 1249-1252, doi:10.1038/nbt1020 (2004).

8 Andrianantoandro, E., Basu, S., Karig, D. K. & Weiss, R. Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* **2**, doi:ARTN 2006.0028 10.1038/msb4100073 (2006).

9 Purnick, P. E. M. & Weiss, R. The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Bio* **10**, 410-422, doi:10.1038/nrm2698 (2009).

10 Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in Escherichia coli. *Nature* **403**, 339-342, doi:10.1038/35002131 (2000).

11 Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335-338, doi:10.1038/35002125 (2000).

12 Atkinson, M. R., Savageau, M. A., Myers, J. T. & Ninfa, A. J. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli. *Cell* **113**, 597-607 (2003).

13 Stricker, J. *et al.* A fast, robust and tunable synthetic gene oscillator. *Nature* **456**, 516-519, doi:10.1038/nature07389 (2008).

14 Tigges, M., Marquez-Lago, T. T., Stelling, J. & Fussenegger, M. A tunable synthetic mammalian oscillator. *Nature* **457**, 309-312, doi:10.1038/nature07616 (2009).

15 Ham, T. S., Lee, S. K., Keasling, J. D. & Arkin, A. P. A tightly regulated inducible expression system utilizing the fim inversion recombination switch. *Biotechnol Bioeng* **94**, 1-4, doi:10.1002/bit.20916 (2006).

16 Siuti, P., Yazbek, J. & Lu, T. K. Synthetic circuits integrating logic and memory in living cells. *Nat Biotechnol* **31**, 448-452, doi:10.1038/nbt.2510 (2013).

17 Bonnet, J., Yin, P., Ortiz, M. E., Subsoontorn, P. & Endy, D. Amplifying genetic logic gates. *Science* **340**, 599-603, doi:10.1126/science.1232758 (2013).

18 Weinberg, B. H. *et al.* Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nat Biotechnol* **35**, 453-462, doi:10.1038/nbt.3805 (2017).

19 Friedland, A. E. *et al.* Synthetic gene networks that count. *Science* **324**, 1199-1202, doi:10.1126/science.1172005 (2009).

20 Anderson, J. C., Voigt, C. A. & Arkin, A. P. Environmental signal integration by a modular AND gate. *Mol Syst Biol* **3**, 133, doi:10.1038/msb4100173 (2007).

21 Nielsen, A. A. *et al.* Genetic circuit design automation. *Science* **352**, aac7341, doi:10.1126/science.aac7341 (2016).

22 Tamsir, A., Tabor, J. J. & Voigt, C. A. Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. *Nature* **469**, 212-215, doi:10.1038/nature09565 (2011).

23 Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C. & Voigt, C. A. Genetic programs constructed from layered logic gates in single cells. *Nature* **491**, 249-253, doi:10.1038/nature11516 (2012).

24 Auslander, S., Auslander, D., Muller, M., Wieland, M. & Fussenegger, M. Programmable single-cell mammalian biocomputers. *Nature* **487**, 123-127, doi:10.1038/nature11149 (2012).

25 Rackham, O. & Chin, J. W. Cellular logic with orthogonal ribosomes. *J Am Chem Soc* **127**, 17584-17585, doi:10.1021/ja055338d (2005).

26 Hooshangi, S., Thiberge, S. & Weiss, R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc Natl Acad Sci U S A* **102**, 3581-3586, doi:10.1073/pnas.0408507102 (2005).

27 Sohka, T. *et al.* An externally tunable bacterial band-pass filter. *Proc Natl Acad Sci U S A* **106**, 10135-10140, doi:10.1073/pnas.0901246106 (2009).

28 Levskaya, A. *et al.* Synthetic biology: engineering Escherichia coli to see light. *Nature* **438**, 441-442, doi:10.1038/nature04405 (2005).

29 Tabor, J. J. *et al.* A synthetic genetic edge detection program. *Cell* **137**, 1272-1281, doi:10.1016/j.cell.2009.04.048 (2009).

30 Weber, W. *et al.* A synthetic time-delay circuit in mammalian cells and mice. *Proc Natl Acad Sci U S A* **104**, 2643-2648, doi:10.1073/pnas.0606398104 (2007).

31 Bayer, T. S. & Smolke, C. D. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nat Biotechnol* **23**, 337-343, doi:10.1038/nbt1069 (2005).

32 Win, M. N. & Smolke, C. D. Higher-order cellular information processing with synthetic RNA devices. *Science* **322**, 456-460, doi:10.1126/science.1160311 (2008).

33 Deans, T. L., Cantor, C. R. & Collins, J. J. A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. *Cell* **130**, 363-372, doi:10.1016/j.cell.2007.05.045 (2007).

34 Rinaudo, K. *et al.* A universal RNAi-based logic evaluator that operates in mammalian cells. *Nat Biotechnol* **25**, 795-801, doi:10.1038/nbt1307 (2007).

35 Winkler, W. C. & Breaker, R. R. Regulation of bacterial gene expression by riboswitches. *Annu Rev Microbiol* **59**, 487-517, doi:10.1146/annurev.micro.59.030804.121336 (2005).

36 Winkler, W., Nahvi, A. & Breaker, R. R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952-956, doi:10.1038/nature01145 (2002).

37 Auslander, S. *et al.* A general design strategy for protein-responsive riboswitches in mammalian cells. *Nat Methods* **11**, 1154-1160, doi:10.1038/nmeth.3136 (2014).

38 Mellin, J. R. *et al.* Riboswitches. Sequestration of a two-component response regulator by a riboswitch-regulated noncoding RNA. *Science* **345**, 940-943, doi:10.1126/science.1255083 (2014).

39 Ren, A., Rajashankar, K. R. & Patel, D. J. Fluoride ion encapsulation by Mg2+ ions and phosphates in a fluoride riboswitch. *Nature* **486**, 85-89, doi:10.1038/nature11152 (2012).

40 Montange, R. K. & Batey, R. T. Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature* **441**, 1172-1175, doi:10.1038/nature04819 (2006).

41 Thore, S., Leibundgut, M. & Ban, N. Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science* **312**, 1208-1211, doi:10.1126/science.1128451 (2006).

42 Frieda, K. L. & Block, S. M. Direct observation of cotranscriptional folding in an adenine riboswitch. *Science* **338**, 397-400, doi:10.1126/science.1225722 (2012).

43 Mandal, M. *et al.* A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science* **306**, 275-279, doi:10.1126/science.1100829 (2004).

44 Serganov, A., Polonskaia, A., Phan, A. T., Breaker, R. R. & Patel, D. J. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **441**, 1167-1171, doi:10.1038/nature04740 (2006).

45 Batey, R. T., Gilbert, S. D. & Montange, R. K. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature* **432**, 411-415, doi:10.1038/nature03037 (2004).

46 Yen, L. *et al.* Exogenous control of mammalian gene expression through modulation of RNA self-cleavage. *Nature* **431**, 471-476, doi:10.1038/nature02844 (2004).

47 Isaacs, F. J. *et al.* Engineered riboregulators enable post-transcriptional control of gene expression. *Nat Biotechnol* **22**, 841-847, doi:10.1038/nbt986 (2004).

48 Rackham, O. & Chin, J. W. A network of orthogonal ribosome x mRNA pairs. *Nat Chem Biol* **1**, 159-166, doi:10.1038/nchembio719 (2005).

49 Meyer, S., Chappell, J., Sankar, S., Chew, R. & Lucks, J. B. Improving fold activation of small transcription activating RNAs (STARs) with rational RNA engineering strategies. *Biotechnol Bioeng* **113**, 216-225, doi:10.1002/bit.25693 (2016).

50 Buskirk, A. R., Kehayova, P. D., Landrigan, A. & Liu, D. R. In vivo evolution of an RNA-based transcriptional activator. *Chem Biol* **10**, 533-540 (2003).

51 Saha, S., Ansari, A. Z., Jarrell, K. A. & Ptashne, M. RNA sequences that work as transcriptional activating regions. *Nucleic Acids Res* **31**, 1565-1570, doi:10.1093/nar/gkg227 (2003).

52 Dueber, J. E., Yeh, B. J., Chak, K. & Lim, W. A. Reprogramming control of an allosteric signaling switch through modular recombination. *Science* **301**, 1904-1908, doi:10.1126/science.1085945 (2003).

53 Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. Computational design of receptor and sensor proteins with novel functions. *Nature* **423**, 185-190, doi:10.1038/nature01556 (2003).

54 Gaber, R. *et al.* Designable DNA-binding domains enable construction of logic circuits in mammalian cells. *Nat Chem Biol* **10**, 203-208, doi:10.1038/nchembio.1433 (2014).

55 Lebar, T., Verbic, A., Ljubetic, A. & Jerala, R. Polarized displacement by transcription activator-like effectors for regulatory circuits. *Nat Chem Biol* **15**, 80-87, doi:10.1038/s41589-018-0163-8 (2019).

56 Fink, T. *et al.* Design of fast proteolysis-based signaling and logic circuits in mammalian cells. *Nat Chem Biol* **15**, 115-122, doi:10.1038/s41589-018-0181-6 (2019).

57 Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-89, doi:10.1038/nature06450 (2008).

58 Ro, D. K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940-943, doi:10.1038/nature04640 (2006).

59 Steen, E. J. *et al.* Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559-562, doi:10.1038/nature08721 (2010).

60 Bayer, T. S. *et al.* Synthesis of methyl halides from biomass using engineered microbes. *J Am Chem Soc* **131**, 6508-6515, doi:10.1021/ja809461u (2009).

61 Ma, S. M. *et al.* Complete reconstitution of a highly reducing iterative polyketide synthase. *Science* **326**, 589-592, doi:10.1126/science.1175602 (2009).

62 Basu, S., Mehreja, R., Thiberge, S., Chen, M. T. & Weiss, R. Spatiotemporal control of gene expression with pulse-generating networks. *Proc Natl Acad Sci U S A* **101**, 6355-6360, doi:10.1073/pnas.0307571101 (2004).

63 Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H. & Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130-1134, doi:10.1038/nature03461 (2005).

64 You, L., Cox, R. S., 3rd, Weiss, R. & Arnold, F. H. Programmed population control by cell-cell communication and regulated killing. *Nature* **428**, 868-871, doi:10.1038/nature02491 (2004).

65 Muller, M. *et al.* Designed cell consortia as fragrance-programmable analog-to-digital converters. *Nat Chem Biol* **13**, 309-316, doi:10.1038/nchembio.2281 (2017).

66 Auslander, D. *et al.* Programmable full-adder computations in communicating three-dimensional cell cultures. *Nat Methods* **15**, 57-60, doi:10.1038/nmeth.4505 (2018).

67 Good, L. Translation repression by antisense sequences. *Cell Mol Life Sci* **60**, 854-861, doi:10.1007/s00018-003-3045-4 (2003).

68 Wagner, E. G. & Simons, R. W. Antisense RNA control in bacteria, phages, and plasmids. *Annu Rev Microbiol* **48**, 713-742, doi:10.1146/annurev.mi.48.100194.003433 (1994).

69 Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-811, doi:10.1038/35888 (1998).

70 Novina, C. D. & Sharp, P. A. The RNAi revolution. *Nature* **430**, 161-164, doi:10.1038/430161a (2004).

71 Zamore, P. D. & Haley, B. Ribo-gnome: the big world of small RNAs. *Science* **309**, 1519-1524, doi:10.1126/science.1111444 (2005).

72 Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science* **294**, 858-862, doi:10.1126/science.1065062 (2001).

73 Lee, R. C. & Ambros, V. An extensive class of small RNAs in Caenorhabditis elegans. *Science* **294**, 862-864, doi:10.1126/science.1065329 (2001).

74 Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853-858, doi:10.1126/science.1064921 (2001).

75 Majdalani, N., Vanderpool, C. K. & Gottesman, S. Bacterial small RNA regulators. *Crit Rev Biochem Mol Biol* **40**, 93-113, doi:10.1080/10409230590918702 (2005).

76 Majdalani, N., Hernandez, D. & Gottesman, S. Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. *Mol Microbiol* **46**, 813-826 (2002).

77 Kruger, K. *et al.* Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **31**, 147-157 (1982).

78 Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849-857 (1983).

79 Doudna, J. A. & Cech, T. R. The chemical repertoire of natural ribozymes. *Nature* **418**, 222-228, doi:10.1038/418222a (2002).

80 Brantl, S. & Wagner, E. G. Antisense RNA-mediated transcriptional attenuation: an in vitro study of plasmid pT181. *Mol Microbiol* **35**, 1469-1482 (2000).

81 Franch, T., Petersen, M., Wagner, E. G., Jacobsen, J. P. & Gerdes, K. Antisense RNA regulation in prokaryotes: rapid RNA/RNA interaction facilitated by a general U-turn loop structure. *J Mol Biol* **294**, 1115-1125, doi:10.1006/jmbi.1999.3306 (1999).

82 Takahashi, M. K. & Lucks, J. B. A modular strategy for engineering orthogonal chimeric RNA transcription regulators. *Nucleic Acids Res* **41**, 7577-7588, doi:10.1093/nar/gkt452 (2013).

83 Westbrook, A. M. & Lucks, J. B. Achieving large dynamic range control of gene expression with a compact RNA transcription-translation regulator. *Nucleic Acids Res* **45**, 5614-5624, doi:10.1093/nar/gkx215 (2017).

84 Green, A. A., Silver, P. A., Collins, J. J. & Yin, P. Toehold switches: de-novo-designed regulators of gene expression. *Cell* **159**, 925-939, doi:10.1016/j.cell.2014.10.002 (2014).

85 Green, A. A. *et al.* Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117-121, doi:10.1038/nature23271 (2017).

86 Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nat Biotechnol* **19**, 751-755, doi:10.1038/90802 (2001).

87 Pardee, K. *et al.* Paper-based synthetic gene networks. *Cell* **159**, 940-954, doi:10.1016/j.cell.2014.10.004 (2014).

88 Slomovic, S., Pardee, K. & Collins, J. J. Synthetic biology devices for in vitro and in vivo diagnostics. *Proc Natl Acad Sci U S A* **112**, 14429-14435, doi:10.1073/pnas.1508521112 (2015).

89 Pardee, K. *et al.* Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell* **165**, 1255-1266, doi:10.1016/j.cell.2016.04.059 (2016).

90 Cordray, M. S. & Richards-Kortum, R. R. Emerging nucleic acid-based tests for point-of-care detection of malaria. *Am J Trop Med Hyg* **87**, 223-230, doi:10.4269/ajtmh.2012.11-0685 (2012).

91 Vinjé, J. Advances in Laboratory Methods for Detection and Typing of Norovirus. *Journal of Clinical Microbiology* **53**, 373-381, doi:10.1128/jcm.01535-14 (2015).

92 Bartsch, S. M., Lopman, B. A., Ozawa, S., Hall, A. J. & Lee, B. Y. Global Economic Burden of Norovirus Gastroenteritis. *PLoS ONE* **11**, e0151219, doi:10.1371/journal.pone.0151219 (2016).

93 Siebenga, J. J. *et al.* Norovirus Illness Is a Global Problem: Emergence and Spread of Norovirus GII.4 Variants, 2001–2007. *The Journal of Infectious Diseases* **200**, 802-812, doi:10.1086/605127 (2009).

94 Elaine, S. *et al.* Foodborne Illness Acquired in the United States—Major Pathogens. *Emerging Infectious Disease journal* **17**, 7, doi:10.3201/eid1701.P11101 (2011).

95 Aron, J. H. *et al.* Norovirus Disease in the United States. *Emerging Infectious Disease journal* **19**, 1198, doi:10.3201/eid1908.130465 (2013).

96 Payne, D. C. *et al.* Norovirus and medically attended gastroenteritis in U.S. children. *The New England journal of medicine* **368**, 1121-1130, doi:10.1056/NEJMsa1206589 (2013).

97 Ahmed, S. M. *et al.* Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis. *Lancet Infect Dis* **14**, 725-730, doi:10.1016/S1473-3099(14)70767-4 (2014).

98 Lopman, B. A., Steele, D., Kirkwood, C. D. & Parashar, U. D. The Vast and Varied Global Burden of Norovirus: Prospects for Prevention and Control. *PLoS Med* **13**, e1001999, doi:10.1371/journal.pmed.1001999 (2016).

99 Kroneman, A. *et al.* Proposal for a unified norovirus nomenclature and genotyping. *Arch Virol* **158**, 2059-2068, doi:10.1007/s00705-013-1708-5 (2013).

100 de Graaf, M., van Beek, J. & Koopmans, M. P. Human norovirus transmission and evolution in a changing world. *Nat Rev Microbiol* **14**, 421-433, doi:10.1038/nrmicro.2016.48 (2016).

101 Ambert-Balay, K. & Pothier, P. Evaluation of 4 immunochromatographic tests for rapid detection of norovirus in faecal samples. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **56**, 194-198, doi:10.1016/j.jcv.2012.11.001 (2013).

102 Vyas, K., Atkinson, C., Clark, D. A. & Irish, D. Comparison of five commercially available immunochromatographic tests for the detection of norovirus in faecal specimens. *The Journal of hospital infection* **91**, 176-178, doi:10.1016/j.jhin.2015.06.013 (2015).

103 Henningsson, A. J. *et al.* Rapid diagnosis of acute norovirus-associated gastroenteritis: evaluation of the Xpert Norovirus assay and its implementation as a 24/7 service in three hospitals in Jonkoping County, Sweden. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* **36**, 1867-1871, doi:10.1007/s10096-017-3005-9 (2017).

104 Shah, M., Chihota, V., Coetzee, G., Churchyard, G. & Dorman, S. E. Comparison of laboratory costs of rapid molecular tests and conventional diagnostics for detection of tuberculosis and drug-resistant tuberculosis in South Africa. *BMC Infectious Diseases* **13**, 352, doi:10.1186/1471-2334-13-352 (2013).

105 Puri, L., Oghor, C., Denkinger, C. M. & Pai, M. Xpert MTB/RIF for tuberculosis testing: access and price in highly privatised health markets. *The Lancet Global Health* **4**, e94-e95, doi:10.1016/S2214-109X(15)00269-7 (2016).

106 Moore, C. *et al.* Evaluation of a broadly reactive nucleic acid sequence based amplification assay for the detection of noroviruses in faecal material. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **29**, 290-296, doi:10.1016/s1386-6532(03)00170-7 (2004).

107 Fukuda, S., Takao, S., Kuwayama, M., Shimazu, Y. & Miyazaki, K. Rapid Detection of Norovirus from Fecal Specimens by Real-Time Reverse Transcription-Loop-Mediated Isothermal Amplification Assay. *Journal of Clinical Microbiology* **44**, 1376-1381, doi:10.1128/JCM.44.4.1376-1381.2006 (2006).

108 Patterson, S. S. *et al.* A nucleic acid sequence-based amplification assay for real-time detection of norovirus genogroup II. *Journal of applied microbiology* **101**, 956-963, doi:10.1111/j.1365-2672.2006.02934.x (2006).

109 Rutjes, S. A., van den Berg, H. H., Lodder, W. J. & de Roda Husman, A. M. Real-time detection of noroviruses in surface water by use of a broadly reactive nucleic acid sequence-based amplification assay. *Appl Environ Microbiol* **72**, 5349-5358, doi:10.1128/aem.00751-06 (2006).

110 Fukuda, S., Sasaki, Y. & Seno, M. Rapid and Sensitive Detection of Norovirus Genomes in Oysters by a Two-Step Isothermal Amplification Assay System Combining Nucleic Acid Sequence-Based Amplification and Reverse Transcription-Loop-Mediated Isothermal Amplification Assays. *Applied and Environmental Microbiology* **74**, 3912-3914, doi:10.1128/AEM.00127-08 (2008).

111 Iturriza-Gomara, M., Xerry, J., Gallimore, C. I., Dockery, C. & Gray, J. Evaluation of the Loopamp (loop-mediated isothermal amplification) kit for detecting Norovirus RNA in

faecal samples. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* **42**, 389-393, doi:10.1016/j.jcv.2008.02.012 (2008).

112 Lamhoujeb, S., Charest, H., Fliss, I., Ngazoa, S. & Jean, J. Real-time molecular beacon NASBA for rapid and sensitive detection of norovirus GII in clinical samples. *Canadian journal of microbiology* **55**, 1375-1380, doi:10.1139/w09-105 (2009).

113 Yoda, T. *et al.* Application of a modified loop-mediated isothermal amplification kit for detecting Norovirus genogroups I and II. *Journal of medical virology* **81**, 2072-2078, doi:10.1002/jmv.21626 (2009).

114 Yaren, O. *et al.* A norovirus detection architecture based on isothermal amplification and expanded genetic systems. *Journal of virological methods* **237**, 64-71, doi:10.1016/j.jviromet.2016.08.012 (2016).

115 Jeon, S. B., Seo, D. J., Oh, H., Kingsley, D. H. & Choi, C. Development of one-step reverse transcription loop-mediated isothermal amplification for norovirus detection in oysters. *Food Control* **73**, 1002-1009, doi:https://doi.org/10.1016/j.foodcont.2016.10.005 (2017).

116 Moore, M. D. & Jaykus, L.-A. Development of a Recombinase Polymerase Amplification Assay for Detection of Epidemic Human Noroviruses. *Scientific Reports* **7**, 40244, doi:10.1038/srep40244 https://www.nature.com/articles/srep40244 - supplementary-information (2017).

117 Guatelli, J. C. *et al.* Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc Natl Acad Sci U S A* **87**, 7797 (1990).

118 Notomi, T. *et al.* Loop-mediated isothermal amplification of DNA. *Nucleic Acids Research* **28**, e63-e63 (2000).

119 Piepenburg, O., Williams, C. H., Stemple, D. L. & Armes, N. A. DNA Detection Using Recombination Proteins. *PLOS Biology* **4**, e204, doi:10.1371/journal.pbio.0040204 (2006).

120 Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-345, doi:10.1038/nmeth.1318 (2009).

121 Green, A. A. Construction and In Vivo Testing of Prokaryotic Riboregulators. *Methods in molecular biology (Clifton, N.J.)* **1632**, 285-302, doi:10.1007/978-1-4939-7138-1_19 (2017).

122 Gupta, N. *et al.* Cross-Reactive Synbody Affinity Ligands for Capturing Diverse Noroviruses. *Anal Chem* **89**, 7174-7181, doi:10.1021/acs.analchem.7b01337 (2017).

123 Zadeh, J. N. *et al.* NUPACK: Analysis and design of nucleic acid systems. *Journal of computational chemistry* **32**, 170-173, doi:10.1002/jcc.21596 (2011).

124 Zadeh, J. N., Wolfe, B. R. & Pierce, N. A. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of computational chemistry* **32**, 439-452, doi:10.1002/jcc.21633 (2011).

125 Furuya, D. *et al.* Age, viral copy number, and immunosuppressive therapy affect the duration of norovirus RNA excretion in inpatients diagnosed with norovirus infection. *Japanese journal of infectious diseases* **64**, 104-108 (2011).

126 Greene, S. R. *et al.* Evaluation of the NucliSens Basic Kit assay for detection of Norwalk virus RNA in stool specimens. *Journal of virological methods* **108**, 123-131 (2003).

127 Gupta, N. *et al.* Whole-Virus Screening to Develop Synbodies for the Influenza Virus. *Bioconjugate Chemistry* **27**, 2505-2512, doi:10.1021/acs.bioconjchem.6b00447 (2016).

128 Rabinowitz, J. A., Lainson, J. C., Johnston, S. A. & Diehnelt, C. W. Non-natural amino acid peptide microarrays to discover Ebola virus glycoprotein ligands. *Chemical Communications*, doi:10.1039/C7CC08242H (2018).

129 Musso, D. *et al.* Molecular detection of Zika virus in blood and RNA load determination during the French Polynesian outbreak. *Journal of medical virology* **89**, 1505-1510, doi:10.1002/jmv.24735 (2017).

130 Waggoner, J. J. *et al.* Viremia and Clinical Presentation in Nicaraguan Patients Infected With Zika Virus, Chikungunya Virus, and Dengue Virus. *Clinical Infectious Diseases* **63**, 1584-1590, doi:10.1093/cid/ciw589 (2016).

131 Khamrin, P. *et al.* Evaluation of immunochromatography tests for detection of novel GII.17 norovirus in stool samples. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* **20** (2015).

132 Ng, A. H. C., Choi, K., Luoma, R. P., Robinson, J. M. & Wheeler, A. R. Digital Microfluidic Magnetic Separation for Particle-Based Immunoassays. *Analytical Chemistry* **84**, 8805-8812, doi:10.1021/ac3020627 (2012).

133 Khalil, A. S. & Collins, J. J. Synthetic biology: applications come of age. *Nat Rev Genet* **11**, 367-379, doi:10.1038/nrg2775 (2010).

134 Friedland, A. E. *et al.* Synthetic Gene Networks That Count. *Science* **324**, 1199-1202, doi:10.1126/science.1172005 (2009).

135 Auslander, S., Auslander, D., Muller, M., Wieland, M. & Fussenegger, M. Programmable single-cell mammalian biocomputers. *Nature* **487**, 123-+, doi:10.1038/nature11149 (2012).

136 Siuti, P., Yazbek, J. & Lu, T. K. Synthetic circuits integrating logic and memory in living cells. *Nat Biotechnol* **31**, 448-+, doi:10.1038/nbt.2510 (2013).

137 Win, M. N. & Smolke, C. D. Higher-order cellular information processing with synthetic RNA devices. *Science* **322**, 456-460, doi:10.1126/science.1160311 (2008).

138 Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. Multi-Input RNAi-Based Logic Circuit for Identification of Specific Cancer Cells. *Science* **333**, 1307-1311, doi:10.1126/science.1205527 (2011).

139 Daniel, R., Rubens, J. R., Sarpeshkar, R. & Lu, T. K. Synthetic analog computation in living cells. *Nature* **497**, 619-+, doi:10.1038/nature12148 (2013).

140 Brantl, S. & Wagner, E. G. H. Antisense RNA-mediated transcriptional attenuation: an in vitro study of plasmid pT181. *Mol Microbiol* **35**, 1469-1482, doi:DOI 10.1046/j.1365-2958.2000.01813.x (2000).

141 Gultyaev, A. P., Franch, T. & Gerdes, K. Programmed cell death by hok/sok of plasmid R1: Coupled nucleotide covariations reveal a phylogenetically conserved folding pathway

in the hok family of mRNAs. *J Mol Biol* **273**, 26-37, doi:DOI 10.1006/jmbi.1997.1295 (1997).

142 Winkler, W., Nahvi, A. & Breaker, R. R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952-956, doi:10.1038/nature01145 (2002).

143 Callura, J. M., Dwyer, D. J., Isaacs, F. J., Cantor, C. R. & Collins, J. J. Tracking, tuning, and terminating microbial physiology using synthetic riboregulators. *P Natl Acad Sci USA* **107**, 15898-15903, doi:10.1073/pnas.1009747107 (2010).

144 Callura, J. M., Cantor, C. R. & Collins, J. J. Genetic switchboard for synthetic biology applications. *P Natl Acad Sci USA* **109**, 5850-5855, doi:10.1073/pnas.1203808109 (2012).

145 Isaacs, F. J. *et al.* Engineered riboregulators enable post-transcriptional control of gene expression. *Nat Biotechnol* **22**, 841-847, doi:10.1038/nbt986 (2004).

146 Lucks, J. B., Qi, L., Mutalik, V. K., Wang, D. & Arkin, A. P. Versatile RNA-sensing transcriptional regulators for engineering genetic networks. *P Natl Acad Sci USA* **108**, 8617-8622, doi:10.1073/pnas.1015741108 (2011).

147 Mutalik, V. K., Qi, L., Guimaraes, J. C., Lucks, J. B. & Arkin, A. P. Rationally designed families of orthogonal RNA regulators of translation. *Nat Chem Biol* **8**, 447-454, doi:10.1038/Nchembio.919 (2012).

148 Rodrigo, G., Landrain, T. E. & Jaramillo, A. De novo automated design of small RNA circuits for engineering synthetic riboregulation in living cells. *P Natl Acad Sci USA* **109**, 15271-15276, doi:10.1073/pnas.1203831109 (2012).

149 Tyagi, S. & Kramer, F. R. Molecular beacons: Probes that fluoresce upon hybridization. *Nat Biotechnol* **14**, 303-308, doi:DOI 10.1038/nbt0396-303 (1996).

150 Zhang, D. Y., Chen, S. X. & Yin, P. Optimizing the specificity of nucleic acid hybridization. *Nat Chem* **4**, 208-214, doi:10.1038/Nchem.1246 (2012).

151 Mutalik, V. K., Qi, L., Guimaraes, J. C., Lucks, J. B. & Arkin, A. P. Rationally designed families of orthogonal RNA regulators of translation. *Nat Chem Biol* **8**, 447-454, doi:10.1038/nchembio.919 (2012).

152 Nissim, L. *et al.* Synthetic RNA-Based Immunomodulatory Gene Circuits for Cancer Immunotherapy. *Cell* **171**, 1138-1150 e1115, doi:10.1016/j.cell.2017.09.049 (2017).

153 Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science* **333**, 1307-1311, doi:10.1126/science.1205527 (2011).

154 Curtis, K. A., Rudolph, D. L. & Owen, S. M. Rapid detection of HIV-1 by reverse-transcription, loop-mediated isothermal amplification (RT-LAMP). *J Virol Methods* **151**, 264-270, doi:10.1016/j.jviromet.2008.04.011 (2008).

155 Hemelaar, J. The origin and diversity of the HIV-1 pandemic. *Trends Mol Med* **18**, 182-192, doi:10.1016/j.molmed.2011.12.001 (2012).

APPENDIX A
COPYRIGHT PERMISSION FOR ADAPTIONS OF FIGURES

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

May 30, 2019

This Agreement between BIODESIGN INSTITUTE –– DUO MA ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4598581306676
License date	May 30, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature
Licensed Content Title	Construction of a genetic toggle switch in Escherichia coli
Licensed Content Author	Timothy S. Gardner et al
Licensed Content Date	Jan 20, 2000
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	Graduate student
Institution name	Biodesign Institute
Expected presentation date	Jul 2019
Portions	Figure 1 Figure 4
Requestor Location	BIODESIGN INSTITUTE 727 E. Tyler St. TEMPE, AZ 85287 United States Attn: BIODESIGN INSTITUTE
Total	0.00 USD
Terms and Conditions	

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

May 30, 2019

This Agreement between BIODESIGN INSTITUTE -- DUO MA ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4598590024068
License date	May 30, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature
Licensed Content Title	A synthetic oscillatory network of transcriptional regulators
Licensed Content Author	Michael B. Elowitz et al
Licensed Content Date	Jan 20, 2000
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	Graduate student
Institution name	Biodesign Institute
Expected presentation date	Jul 2019
Portions	Figure 1 Figure 2
Requestor Location	BIODESIGN INSTITUTE 727 E. Tyler St. TEMPE, AZ 85287 United States Attn: BIODESIGN INSTITUTE
Total	0.00 USD
Terms and Conditions	

THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE LICENSE TERMS AND CONDITIONS

May 30, 2019

This Agreement between BIODESIGN INSTITUTE -- DUO MA ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science") consists of your license details and the terms and conditions provided by The American Association for the Advancement of Science and Copyright Clearance Center.

License Number	4598590125960
License date	May 30, 2019
Licensed Content Publisher	The American Association for the Advancement of Science
Licensed Content Publication	Science
Licensed Content Title	Synthetic Gene Networks That Count
Licensed Content Author	Ari E. Friedland,Timothy K. Lu,Xiao Wang,David Shi,George Church,James J. Collins
Licensed Content Date	May 29, 2009
Licensed Content Volume	324
Licensed Content Issue	5931
Volume number	324
Issue number	5931
Type of Use	Thesis / Dissertation
Requestor type	Scientist/individual at a research institution
Format	Print and electronic
Portion	Figure
Number of figures/tables	1
Order reference number	
Title of your thesis / dissertation	Graduate student
Expected completion date	Jul 2019
Estimated size(pages)	1
Requestor Location	BIODESIGN INSTITUTE 727 E. Tyler St. TEMPE, AZ 85287 United States Attn: BIODESIGN INSTITUTE
Total	0.00 USD
Terms and Conditions	

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

May 30, 2019

This Agreement between BIODESIGN INSTITUTE -- DUO MA ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4598590248338
License date	May 30, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature
Licensed Content Title	Programmable single-cell mammalian biocomputers
Licensed Content Author	Simon Ausländer, David Ausländer, Marius Müller, Markus Wieland, Martin Fussenegger
Licensed Content Date	Jun 3, 2012
Licensed Content Volume	487
Licensed Content Issue	7405
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	Graduate student
Institution name	Biodesign Institute
Expected presentation date	Jul 2019
Portions	Figure 1
Requestor Location	BIODESIGN INSTITUTE 727 E. Tyler St. TEMPE, AZ 85287 United States Attn: BIODESIGN INSTITUTE

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

May 30, 2019

This Agreement between BIODESIGN INSTITUTE -- DUO MA ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4598590456035
License date	May 30, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature
Licensed Content Title	Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch
Licensed Content Author	Alexander Serganov et al
Licensed Content Date	May 21, 2006
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	Graduate student
Institution name	Biodesign Institute
Expected presentation date	Jul 2019
Portions	Figure 4
Requestor Location	BIODESIGN INSTITUTE 727 E. Tyler St. TEMPE, AZ 85287 United States Attn: BIODESIGN INSTITUTE
Total	0.00 USD
Terms and Conditions	

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

May 30, 2019

This Agreement between BIODESIGN INSTITUTE -- DUO MA ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4598590946375
License date	May 30, 2019
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Biotechnology & Bioengineering
Licensed Content Title	Improving fold activation of small transcription activating RNAs (STARs) with rational RNA engineering strategies
Licensed Content Author	Sarai Meyer, James Chappell, Sitara Sankar, et al
Licensed Content Date	Sep 9, 2015
Licensed Content Volume	113
Licensed Content Issue	1
Licensed Content Pages	10
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 1
Will you be translating?	No
Title of your thesis / dissertation	Graduate student
Expected completion date	Jul 2019
Expected size (number of pages)	1
Requestor Location	BIODESIGN INSTITUTE 727 E. Tyler St. TEMPE, AZ 85287 United States Attn: BIODESIGN INSTITUTE
Publisher Tax ID	EU826007151

ELSEVIER LICENSE
TERMS AND CONDITIONS

May 30, 2019

This Agreement between BIODESIGN INSTITUTE -- DUO MA ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4598591107038
License date	May 30, 2019
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cell
Licensed Content Title	Toehold Switches: De-Novo-Designed Regulators of Gene Expression
Licensed Content Author	Alexander A. Green,Pamela A. Silver,James J. Collins,Peng Yin
Licensed Content Date	Nov 6, 2014
Licensed Content Volume	159
Licensed Content Issue	4
Licensed Content Pages	15
Start Page	925
End Page	939
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Figure 1
Title of your thesis/dissertation	Graduate student
Publisher of new work	Biodesign Institute
Expected completion date	Jul 2019
Estimated size (number of pages)	1
Requestor Location	BIODESIGN INSTITUTE 727 E. Tyler St.

ELSEVIER LICENSE
TERMS AND CONDITIONS

May 30, 2019

This Agreement between BIODESIGN INSTITUTE –– DUO MA ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4598591392097
License date	May 30, 2019
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cell
Licensed Content Title	Paper-Based Synthetic Gene Networks
Licensed Content Author	Keith Pardee,Alexander A. Green,Tom Ferrante,D. Ewen Cameron,Ajay DaleyKeyser,Peng Yin,James J. Collins
Licensed Content Date	Nov 6, 2014
Licensed Content Volume	159
Licensed Content Issue	4
Licensed Content Pages	15
Start Page	940
End Page	954
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Figure 3
Title of your thesis/dissertation	Graduate student
Publisher of new work	Biodesign Institute
Expected completion date	Jul 2019
Estimated size (number of pages)	1
Requestor Location	BIODESIGN INSTITUTE 727 E. Tyler St.