Multi-Ribozyme-gRNA-Aptazyme (RGA) networks for biosensing and gene regulation

Chenggang Xi

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2019

Reading Committee:

James Carothers, Chair

Eric Klavins

Program Authorized to Offer Degree:

Chemical Engineering

University of Washington

Abstract

Multi-Ribozyme-gRNA-Aptazyme(RGA) networks for biosensing and gene regulation

Chenggang Xi

Chair of the Supervisory Committee:
Assistant Professor James Carothers
Chemical Engineering

Biological parts can be considered as molecular building blocks to assemble complex synthetic genetic networks utilizing synthetic biology. In this work, we constructed three types of multi-Ribozyme-gRNA-Aptazyme (RGA) networks for biosensing and gene regulation by linking RGA, Ribozyme-gRNA-Ribozyme (RGR) and CRISPRi systems together. We screened out two different aptazymes that fit this network and analyzed their one-layer and two-layer single-input RGAs behaviors and functions in two-inputs systems with different architectures. Overall, we built two IMPLY gates, two NIMPLY gates, one AND gate, one partial NAND gate, one conditional non-Boolean gate. The RGA shows as much as a 7.38-fold increase in ligand-dependent system-output. The interaction between RGAs that are inserted in circuits in parallel has also been reported, which contributions to the predictability of more complicated RGA networks.

TABLE OF CONTENTS

Chapter1 INTRODUCTION

1.1 Motivation

Cells are dynamical systems that constantly react and evolve within a changing environment. Based on surrounding chemical or biological stimuli, computation can be performed in living cells by DNA- or RNA- encoded circuits that process sensory information and control biological functions. Also, living cells are able to make decisions based on information processing genetic programmes and many of them execute digital functions. With the capability to build programmable gene regulatory circuits in living cells, engineers could build novel decision-making regulatory networks for use in a variety of applications, ranging from gene therapies, tissue engineering, bioproducts production to metabolic engineering.

With the help of synthetic biology, we can program cells to analyze molecular entities and report their environment in real-time. In order to realize gene regulation or biosensing utilizing cells, the most convenient way is to encode ligand-inducible transcription factor systems into them. Cells can can exhibit the correct ON and OFF characteristics in the presence and absence of inducer, thus forming a genetic device with digital input-output signals. Apart from mimicking the logic gate operation, more complex structures and circuits can be constructed in more sophisticated areas. By this work, we want to build genetic circuits that are responsive to multiple small molecules with predictable gene regulation patterns.

1.2 RNA-based platforms

RNA is a dynamic, complex and structured genetic element that is essential for mediating biological activities, capable of catalysis and is used for cellular control schemes. Nowadays, CRISPR-Cas9, based on nuclease inactive Cas9 (dCas9) derived from *Streptococcus pyogene*(Jinek et al., 2012) has been employed commonly for the programmed regulation of transcription and shows promise as a platform for building predictable gene regulatory networks. CRISPR-dCas9 networks are composed of multiple guide RNAs (gRNAs) that, upon binding to dCas9, can inhibit or activate their own transcription or the transcription of specific downstream genes. However, this kind of regulation systems can’t control gene expression quantitatively because they can only achieve a constant ON or OFF state, so most of the CRISPR-dCas9 systems are still used for genome editing purpose rather than tuning the gene expression.

Meanwhile, riboswitches are RNA-based genetic control elements that regulate a wide set of basic metabolic pathways in prokaryotes. The ribozyme-based device framework supports genetic controllers in different organisms, is responsive to diverse ligands, exhibits complex computation, and has been applied to regulate complex phenotypes. Novel Ribozyme-gRNA-Ribozyme (RGR) (Fig. 1 (a)) architectures by encoding cis-cleaving ribozymes (Rbz) that flank the gRNA have been used to drive gRNA expression, whose gRNA site can cooperate with CRISPR-dCas9 system to further regulate the gene expression such as GFP production is this case. Meanwhile, the

level of RGR-inputs can be quantified by β-Estradiol concentration, which will activate the pGalZ4 promoter and then increase the output of RGR element. The flanked ribozymes of RGR would cleave themselves upon being expressed, thus exposing the gRNA binding site to dCas9-Mxi1 and repressing the GFP production (Fig. 1 (b)). By measuring the responses of CRISPRi repression cascades as sensitivities to RGR-inputs, design elements impacting gRNA expression and function from RGR were identified.

1.3 Gene regulatory circuits

Utilizing the tools of synthetic biology, numerous genetic circuits have been implemented such as engineered logic operation in analog and digital circuits. Synthetic gene circuits are constructed by rewiring transcription factors and promoters to create novel regulatory topologies.

Many classes genetic parts exhibit characteristics that are applicable for creating genetic circuits have been established. RGRs have been used as a successful means of Pol II-mediated gRNA expression in mammalian (Nissim et al., 2014) , amphibian (Fei et al., 2016) , fungal(Weber et al., 2017) and yeast(Gander et al., 2017) hosts. Despite their host portability and prevalence, little work has been done to characterize the gRNA expression efficiency from an RGR nor is it fully understood if 5' Rbz cleavage is necessary (Gander et al., 2017; Yoshioka et al., 2016) . Though the RGR architecture has been doubled and tripled to express multiple gRNA from a single transcript (Xu et al., 2017; Yoshioka et al., 2016) no variants have been built with different cis-cleaving ribozymes. Consequently, it is unknown how different Rbz structures or post-cleavage sequence 'scars' impact gRNA expression and function. As the sequence of the gRNA handle (Briner et al., 2014; Mekler et al., 2016) and gRNA target (Chuai et al., 2017; Slaymaker et al., 2015; Wong et al., 2015) have significant impact on gRNA function it is expected that additional 5' or 3' sequence would negatively impact function. Finally, despite successes in appending RNA ligand-sensing aptamers to other gRNA expression systems to create inducibly active gRNA (Liu et al., 2016a; Tang et al., 2017), no work has been done to extend such functionality to RGRs. A quantitative understanding of gRNA expression efficiency from the RGR architecture and an understanding of how different cis-cleaving RNA effect gRNA expression will determine how engineerable it is a gRNA expression platform.

These insights can lead to the creation of the Ribozyme-gRNA-Aptazyme (RGA) architecture. This kind of gRNA regulatory system has been developed by replacing 3' RGR Rbz with one ligand-sensing aptazyme in order to extend functionality of RGRs. Aptazyme is one kind of aptamer-regulated Rbz that would activate or inhibit the ribo-switch once binding to target molecule. In this system, the gRNA expression will not only be affected by the concentration of β-Estradiol but also controlled by the metabolite which the exact aptazyme can detect. Once the aptazyme has been combined with the specific ligand, the steric hindrance would prevent gRNA from binding of dCas9-Mxi1, so it would repress the GFP production any more. With this design, we can create small-molecule inducible gRNA expression system programmed at the level of the gRNA and build even more complicated biological logic circuits with multiple input variables.

1.4 Overview of thesis contribution

This thesis is mainly based on the project developed by Dr. William E. Voje Jr. He has constructed twelve different RGRs and researched the efficacy and competition effects of these modified gRNAs, two one-layer RGA that binding Theophylline (Theo(A)-CAAGUGAA and Theo(A)-AAAAA) integrated architectures has also been builded. However, the RGA pool is so limited that constrain our further research, so the first part of this thesis is to find other potential aptazyems that are effective in our RGA context.

Moreover, we've done in-depth study in installing RGRs in series and get accurate simulation results. Whereas, it's hard to computationally predict the result of the integration that arrange different RGAs in parallel or separate them in different layers. Because we didn't know if the pattern after the combination is a simple addition in gene expression levels or would cause interactions between the RGA parts, and if it has interaction, we also need to know if it's an antagonism or a synergism to identify the degree of disturbance. Herein, I integrated other kinds of RGAs with various underlaying architectures and interpreted their functions within changing chemical environment in the next part.

Above all, we screened four types of RGAs by introducing theophylline (Theo), guanin (Gua), neomycin (Neo), and Tetracycline (Tet) aptamer sequence into the well-developed RGR structure. These aptamers are all known to function in eukaryotic cells with wide dynamic ranges and an ability to get these ligands into the cells. Afterwards, two different aptazymes Theo (Theo(A)-CAAGUGAA) and Tet has been successfully employed in RGA networks with one-layer and two-layer cascades. The response dynamics of these architectures were readily modeled and as much as a 10-fold increases in ligand-dependent system-output were observed. Moreover, three brand-new types of multi-inputs RGA architectures have been designed which allow us to construct digital circuits or sensors of great complexity in biological systems. Finally, we thoroughly analyzed the response of each multi-RGA networks and explained their behavior so as to lays a foundation for future research in multi-input biosensing circuits.

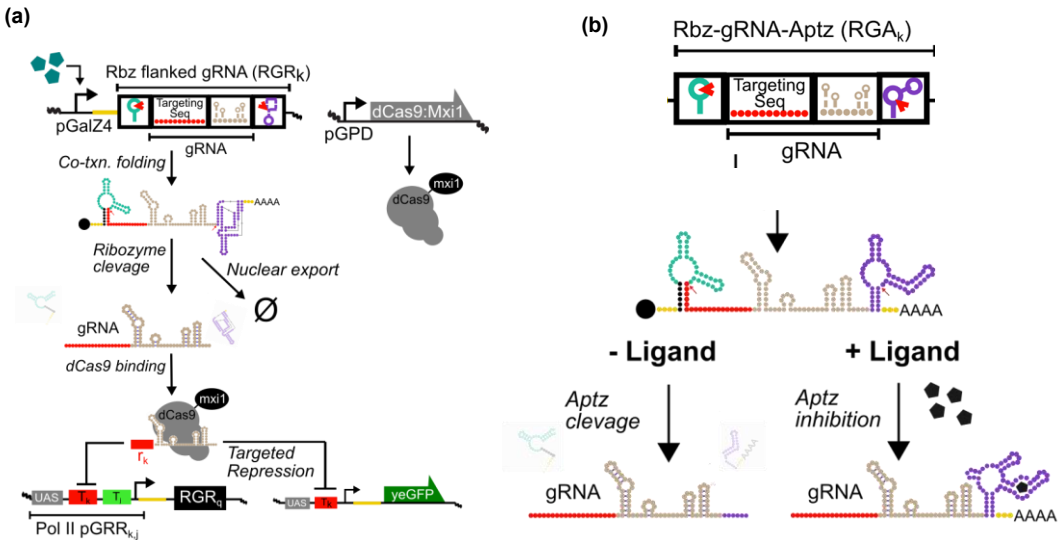


Figure1 (a) The architecture of RGR and its mechanism: with more β-Estradiol induction more GFP gene would be repressed. **(b)** The architecture of RGA and its mechanism: with more ligand, less gRNA would be expressed so less GFP gene would be repressed.

Chapter 2: Aptazyme-regulated gRNA expression with RGAs

2.1 Circuit design

This part of work was accomplished through the construction and analysis of 4 different RGA devices. Based on the former experiences on developing 44 different RGRs, we find that RGR inputs without 5' Rbzs require 1.8-3.6X fold more transcriptional input to achieve the same output as their RGR counterparts. Besides, RGR inputs without 3' Rbzs result in non-functional gRNA expression which means we can show the activity of an aptazyme in the best way by flanking it into the 3' end of the gRNA. We also examine the versatility of the RGR architecture by substituting the original 5' and 3' Rbz with five different hammerhead Rbz variants with diverse sequences and secondary structure topologies. We find that functional RGRs can be engineered from these Rbz and show that post-cleavage scar sequences can cause up to 7.38-fold reductions in gRNA activity (Fig 2.4 (c)). Therefore, we choose 4 aptazymes with response to 4 different ligands to only substitute the 3' end Rbz, while leave the 5' end still be a Rbz. (Fig 2.1)

In order to characterize our experimental yeast systems, we tried to design a fundamental genetic device with ON and OFF state controlled by β-estradiol(bE) inducible pGalZ4 promoter and linked it with our RGA. This promoter can be analogous to a digital switch, but there wouldn't be a logistic 0 to 1 signal because the promoter has continuously change with the transition of bE concentration from low to high. This property of our system enables us to describe the signals with a Hill function and to predict the functions of multi-layer networks based on each single layer activities. Finally, we used the yeast enhanced green fluorescent protein (yeGFP) which is a versatile and measurable reporter of gene expression in *C. albicans* and *Saccharomyces cerevisiae* with a $pGRR_{i,j}$ promoter, which contains multiple binding target of $gRNA_{i,j}$-dCas9-mix1 repressor. Thus, when we increase the bE concentration, pGal promoter would start express RGA output and once the RNA folds properly, 5' Rbz and 3' Aptazyme will automatically cleave themselves which expose the gRNA to dCas9-mix1 protein and then the whole complex would target the pGRR promoter and repress the downstream yeGFP transcription and then cause the decrease of overall fluorescence, thus showing us an ON to OFF switch.

In this work, we find

2.2 Methods

RGA design and plasmid construction

All plasmids constructed were for stable integration into the yeast genome. Plasmids designs were assembled based on the existing Theo(A)-CAAGUGAA RGA plasmid. Recurring genetic elements including promoters, yeast integration makers, ribozymes, gRNA in backbone, while the inserts including guanin, neomycin, Tetracycline aptamer sequences are all come from literature. The Golden Gate method was used for plasmid assembly (Engler, C. et al. 2009). Cloned plasmids were transformed into DH10B E. coli and plated on LB agar plates containing 100 mg/mL carbenicillin and sequence verified (Eurofins). Correct colonies were grown overnight at 37C in LB under 100 mg/mL carbenicillin selective pressure. Plasmid DNA was extracted and purified using

QIAGEN miniprep kit.

Creation of stable yeast strains

Ultimately all strains tested were diploid W303 Saccharomyces cerevisiae strains. All components were first integrated into the haploid background strains MATa W303-1A and MATalpha W303-1B. Stable yeast transformations were done using a lithium acetate protocol162 Plasmids were first linearized with a Pmei (NEB) digestion. Linearised fragments are mixed with necessary reagents and transformed at 42C. Cells were plated on SDO with the appropriate selective pressure. Following this transformation colony PCR were used to confirm that the gene element was successfully integrated. Stable diploid yeast strains were generated by co-culturing MATa and MATalpha strains and subsequently plating the culture onto SDO with two selection markers that were unique to MATa and MATalpha strains.

Experimental yeast growth conditions

Three colonies were picked to inoculate three wells of 96 deep well plate containing 1 mL of SC media. This plate was then incubated overnight at 30C shaking at 800 RPM. The following day wells containing cells were back-diluted 1:1000 in 1mL fresh SC containing the required inducer and were incubated under the same conditions for 16 hours. To best approximate steady state growth, these cells were then diluted 1:50 in 1mL fresh media with appropriate inducers and allowed to grow for an additional 6 hours.

Calibrating experimental data

Calibration curves mapping β-estradiol input-concentrations to transcriptional strength were generated using a diploid yeast strain strain with pGalZ4 driving the expression of yeGFP. For calibration biological triplicates were exposed to six to twelve β-estradiol concentrations and tested using flow-cytometry. Median fluorescent values were background subtracted using a no fluorescence negative control strain. This data fit using a Hill equation.

Testing one-layer RGAs

For one-layer cascades, three biological replicates were completed for each condition. Two β-estradiol series (0.01, 0.5, 1, 3, 8, 20, 100, 500 nM) were completed five Theophylline series (0, 5, 10 ,15, 20, 35 mM), five Tetracycline series (0, 0.5, 1, 1.5, 2 mM), three Neo(0, 0.1, 1 mM) and two Gua (0, 0.2 mM) conditions. Finally, three biological replicates were run for a positive control (diploid strain containing with no RGR expression) and a negative control (diploid strain with no fluorescent reporter) at the same Theophylline series.

Flow cytometry

A BD Accuri C6 flow cytometer equipped with a CSampler plate adapter was used to assay CRISPRi network sfGFP output fluorescence. Excitation wavelengths of 488 and 640 nm and an emission detection filter at 533 nm were used (FL1 channel). 200

uL of cells are transferred from 96 deep well plate to a 96 well assay plate. At least 15,000 cells were measured above a 400,000 AU FSC-H threshold (to screen out cellular debris). These data were then gated for the healthy yeast cells (SSC-A, FSC-A) and the resulting median/mean FL-1 values were calculated using custom python scripts.

Fitting system response with a Hill-equation

The input-output response of one- and two- layer cascades were modeled using a Hill-equation. After conversion of input- and output-responses to TU, data are fit with a Hill equation [f(L, vMin, vMax, n, Ka) = (vMax - vMin)/((Ka)^n + 1) + vMin] where L is the input value, vmin is the minimum expression, vmax is the maximum expression, n is the Hill coefficient, and ka is the input amount resulting in half-saturation of the response.

In vitro cleavage validation

Prepare in vitro transcription template by PCR amplifying previously purified DNA oligo with Phusion High-Fidelity DNA Polymerase. Design an appropriate 30 reverse primer and a 50 forward primer containing a standard T7 promoter sequence. Perform in vitro transcription reactions in PCR tubes. The final reaction volume per sample is 50 µl. Prepare two reactions per construct: one in the presence of saturating ligand concentrations and one without ligand. Include a no-polymerase control. Place capped reaction tubes into a prewarmed thermocycler. Incubate for 5 min without polymerase at 37 C. Add polymerase to the reaction tubes without removing them from the thermocycler. Quickly mix by pipetting up and down, recap and begin timing. At intervals ranging from 5 min to 1 h carefully withdraw 10 µl of the reaction. Stop the reaction by adding it to a PCR tube containing 10 µl of ice-cold 2X loading buffer, mix well and place them on ice. Quenched aliquots can be stored at 20 C or used immediately for PAGE. Cast, load, and run 8% Urea PAGE gel, as in Oligo purification. When loading, include your no-polymerase control sample as reference wich differentiate template DNA bands from the RNA bands. Disassemble apparatus, remove glass plate and carefully transfer gel into Sybr Gold staining solution (10 Sybr Gold in 0.5 TBE buffer). Place the container on an orbital shaker and gently shake (100 rpm), protected from light, for 15 min. Photograph gel on a 400 nm UV transilluminator. Ensure exposure time is sufficient to yield a non-saturated image.

2.3 Experimental Results

2.3.1 BE induction Calibration

We first tested a pGal induced expression control strain of yeGFP. β-estradiol can induce the pGalZ4 promoter producing yeGFP(Figure 2) and the GFP fluorescence can be detected and calculated by flow cytometer, so we can get a fluorescence vs. β-estradiol plot to describe the relationship between β-estradiol concentration and pGal promoter activation. Then we fit the data with Hill function and get a switch-like response, the dynamic range of response to β-estradiol concentration is from 1 to 500 nM so we should design the further experiments within this range of concentration. We

also notice that 500 nm will be the saturation stage of pGal promoter so we can suppose there is only background effect when we titrate other device with β-estradiol concentration higher than 500 nm. This is the basis of our background correct method that we can consider 500 nM bE as a complete OFF state and erase anything larger than control's fluorescence.

2.3.2 RGA screening

We've screened all the 4 different aptazymes(Fig S1 (a-d)) respectively binding different ligands [Fig 2.1(a-d)] for their ability to fold properly, self-cleave, and respond to ligands as expected. We also run the RGR control to show if these ligands would have universe effect on other part of this device. There of them show normal post-cleavage gRNA activities but only two give us ligand-responsive signals. All the devices showing us proper cleavage are those with a relative short post-cleavage scar sequence[pic], which supports our hypothesis that long scar sequence would inhibit gRNA's ability to bind with dCas9 protein, then they won't be able to repress the GFP reporter. Among them, Tet and Theo devices show desired ON to OFF switch, and with the addition of metabolites, the fluorescence would restore back to the ON state. These two one-layer RGA both compose a IMPLY logic gate. However, Neo has proper fold and cleavage activities while Neo has universe inhibition on ribosome cleavage, and our yeast strain includes a kanamycin-resistance gene which provides resistance to Neo, so there is always less fluorescence after we add neo into yeast culture buffer. For other aptamers like Gua, Ka-Theo, and p-AF, they will leave long sequence of post-cleavage scars on the 5' end, so the guide RNA with large tail won't be able to bind with dCas9 effectively and lose the ability to repress the pGRR promoter for GFP. Therefore, we successfully assembled 3 one-layer RGA devices and investigated their ability to modulate gene expression with response to their environment consisting of different metabolites concentrations.

2.3.3 One-layer Theo RGA titration (IMPLY gate)

While we prepare the Theophylline solution, we find its maximum solubility in SC buffer is about 40 mM and needle-like crystal will precipitate after overnight standing, which means 40 mM Theo is an oversaturated solution. Therefore, we use 35 as the max value for Theo RGA tests. Afterwards, we apply an RGR control test and found that 35 nM Theo would cause large background for RGR constructs as well as our RGA device. Then we reduce the concentration again to 20 mM which has background less than 9% in positive control and 6% in the negative control.

Theoretically, adding Theophylline reduces the amount of aptazyme cleavage which in-turn should reduce functional gRNA expression (Figure 3(a)) and we choose the Theo(A)-CAAGUGAA RGA yeast strain which shows larger fold change for later construction. In a general scale, the Theo RGA shows an inducible one-layer repression cascade as the increase of β-estradiol concentration, which has the same behavior as RGR devices (Figure 3(b)). Then as the rise of Theophylline concentration, the GFP fluorescence would increase and gradually come back to a relatively high level (Figure 3(c)) and we can see in low bE induction level, there will be a saturation point where the relative GFP approaches 1. But RGA devices would produce gRNA

with a 3’ scar that will still affect the expression of downstream genes, so adding the Theophylline aptamer won’t totally cancel the influence of β-estradiol induction. Moreover, the output change can be up to 10 folds from the highest 20 mM Theophylline induction to no Theophylline (Figure 2.4 (b)) and when the β-estradiol induction is more significant in high repression regime, the effect of Theophylline will be less dominant and the fold change starts to decrease into 1.

2.3.4 One-layer Tet RGA titration (IMPLY gate)

At the first time, we cultured Tet RGA in sc buffer with 1-5 mM Tet, and we found Tet will be toxic to cells higher than 2 nm (Figure 2.1 (b)). Therefore, we used 1.6-2 mM Tet as the maximum concentration and get the response curve of one-layer Tet RGA. Because Tet generates fluorescence by itself, we apply a background correction by keeping the FL1 median value in 500 nm bE induction level of each Tet titration condition as the same. Also, because the pGal promoter is slightly leaky, the start point in extreme low bE induction level (0.01 nM) the fluorescence would increase a little with the rise of Tet concentration and 0.8 mM is about the saturation value of this effect. Tet buffer higher than 0.8 mM will cause a burden on cells and make the system hard to calibrate like high Theo situation. We can also see the wired transfer function curve of Tet device’s fluorescence over ligand induction value with second and third switching points when ligand concentration is larger than 1 Mm. Therefore, we use 0.75-0.8 mM as the maximum values for later characterization. We also ran an RGR control which shows the Tet solution lower than 1 nM has an identical increase caused by the color of Tet itself and supports our background elimination method. The Tet device still gives us good ligand-responsive inhibition and increases the fluorescence by 2.33X fold even we use much smaller concentration as we did in Theo device. And when we compare these two devices, the Ka of Tet device are much smaller than Theo RGA, which means that Tet aptazyme’s cleavage rate is faster than Theo aptazyme.

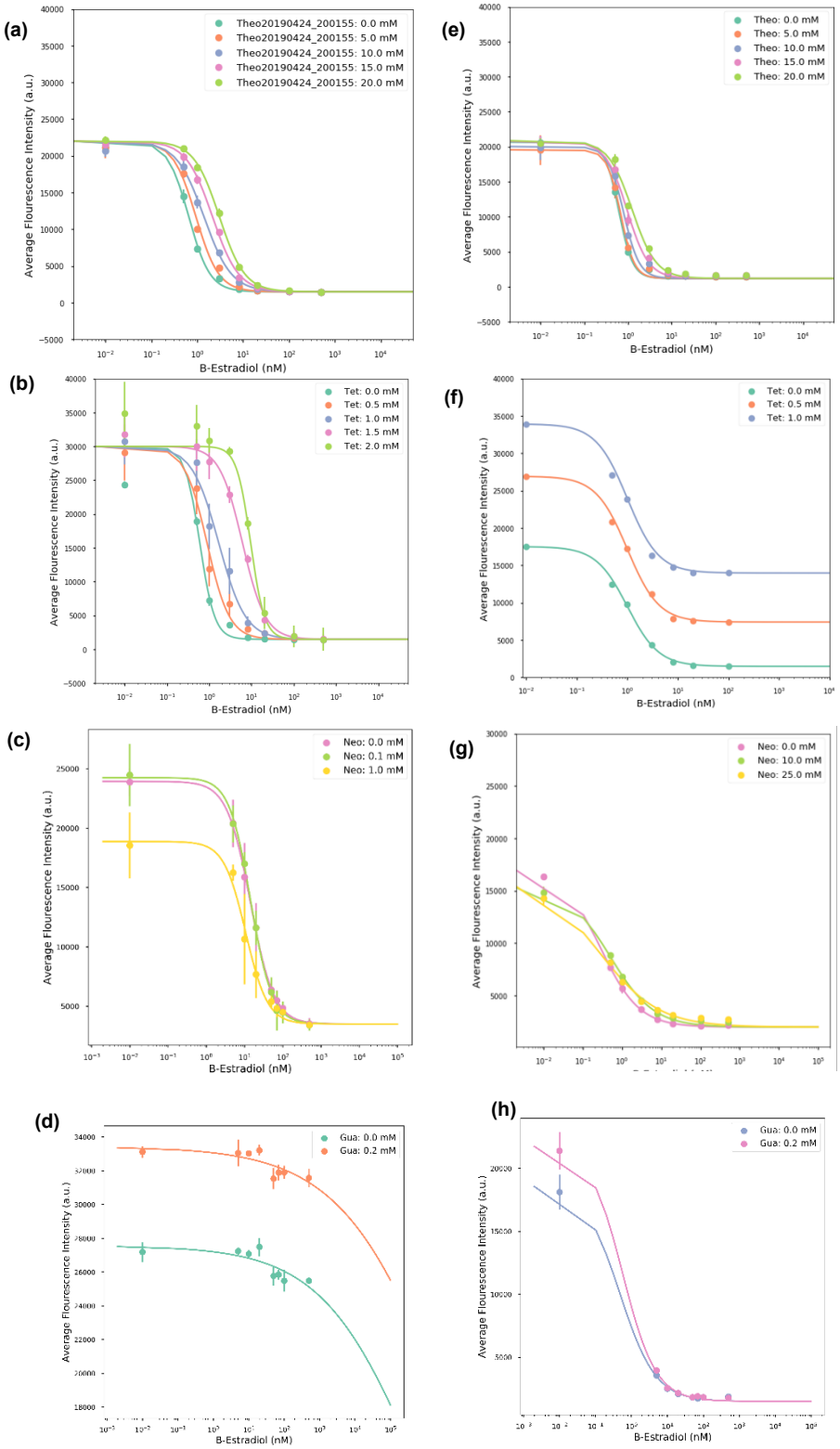


Figure 2.1 The ligand responsive dynamics of different RGA-W8 (a)Theo-RGAW8 (b) Neo-RGAW8 (c) Gua-RGAW8 (d) Tet-RGAW8; The responsive curves of control RGR-w8 within different ligands and concentrations (e)Theophylline (f)Neomycin (g)Guanine (h)Tetracycline

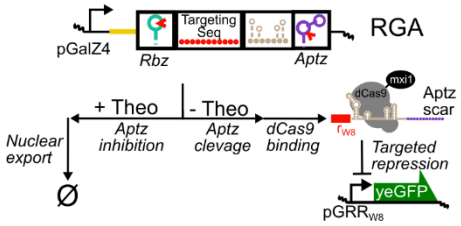


Figure 2.2 Diagram of the of the Rbz-gRNA-aptazyme architecture (RGA). In the absence of theophylline the 3' theophylline aptazyme cleaves, resulting in the functional expression of gRNA. In the presence of theophylline aptazyme cleavage is inhibited, resulting in gRNA export from the nucleus.

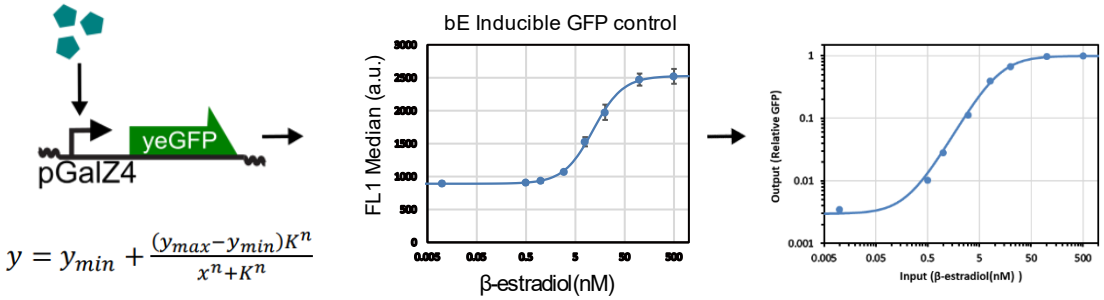
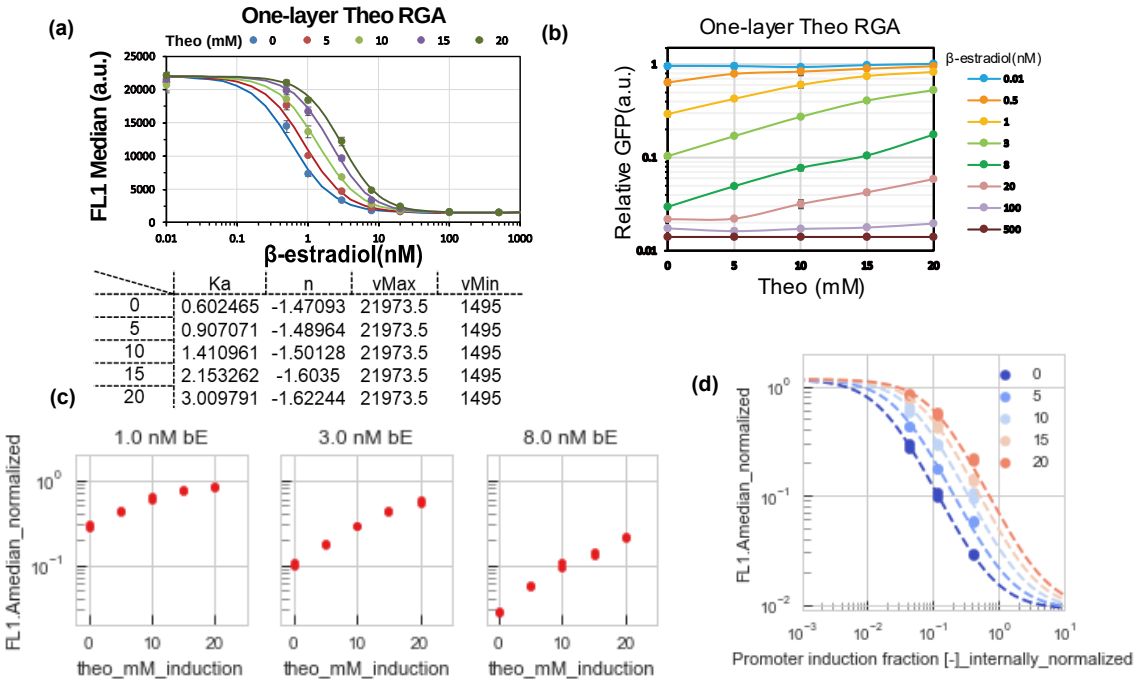
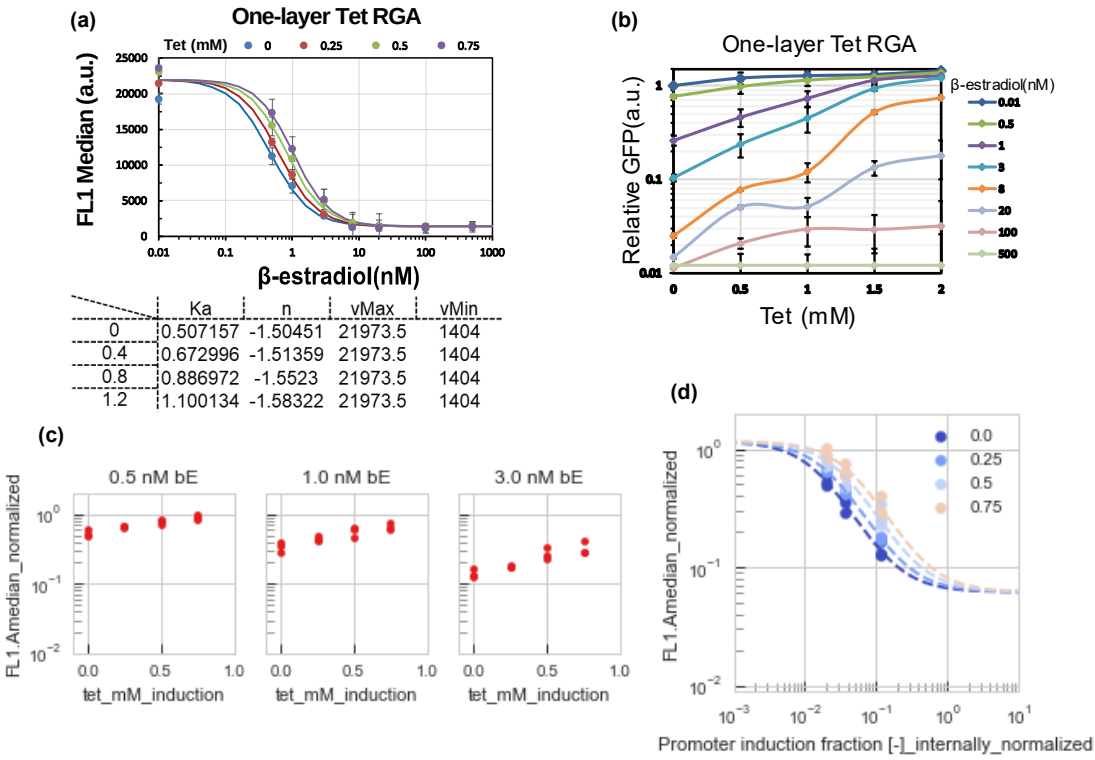


Figure 2.3 The inducible promoter pGalZ4 was calibrated using a fluorescent reporter with a Hill function.



	Ka	n	vMax	vMin
0	0.602465	-1.47093	21973.5	1495
5	0.907071	-1.48964	21973.5	1495
10	1.410961	-1.50128	21973.5	1495
15	2.153262	-1.6035	21973.5	1495
20	3.009791	-1.62244	21973.5	1495

Figure 2.4 (a) The Theophylline and β-estradiol responsive dynamics plot of Theo RGA; the table consists of parameters from Hill equation fitting curves of RGA fluorescence data **(b)**The transfer functions of Theo-RGA in different β-estradiol concentrations **(c)** Biological triplicate titration response data show increasing Theophylline inhibits RGA mediated gRNA repression at multiple levels of induction (increasing bE induction). The fold changes of their relative GFP are 1.0 nM bE: 2.94; 3.0 nM bE: 5.50; 8.0 nM bE: 7.38. **(d)** The Theophylline and β-estradiol responsive dynamics plot of Theo RGA with a relative GFP as output and normalized pGal promoter induction strength as input.



	Ka	n	vMax	vMin
0	0.507157	-1.50451	21973.5	1404
0.4	0.672996	-1.51359	21973.5	1404
0.8	0.886972	-1.5523	21973.5	1404
1.2	1.100134	-1.58322	21973.5	1404

Figure 2.5 **(a)** The Tetracycline and β-estradiol responsive dynamics plot of Tet-RGA; the table consists of parameters from Hill equation fitting curves of RGA fluorescence data **(b)** The transfer functions of Tet-RGA in different β-estradiol concentrations **(c)** Biological triplicate titration response data show increasing Theophylline inhibits RGA mediated gRNA repression at multiple levels of induction (increasing bE induction). The fold changes of their relative GFP are 0.5 nM bE: 1.68; 1.0 nM bE: 1.96; 3.0 nM bE: 2.33. **(d)** The Tetracycline and β-estradiol responsive dynamics plot of Tet-RGA with a relative GFP as output and normalized pGal promoter induction strength as input.

Chapter 3: multi-layer RGA design

RGR principle

3.1 Circuit design

With both RGA and RGR components developed, we are able to integrate the ligand-responsive sensor into the NOR gates in which we've already got in-depth study. The simplest multi-layer logical circuit we can assembly is a single RGA two-layer cascade. We employed pGRR-W36W8-RGR-W20 which is targeted by W8 AND W36 gRNA as input and express W20 gRNA as output, and W20 can further repress the RGP reporter with a pGRR-W17W20 promoter. This RGR can be considered as an invertor that converts high W8 gRNA signal into low W20 gRNA output and makes the entire device an OFF to ON switch with the addition of bE.

Furthermore, with RGA embedded, we can turn the system responsive to small molecule metabolites as well. As discussed in the last part, we could inhibit the aptazyme cleavage ability by the induction of ligand, it will decrease the amount of effective gRNA1 and relieve the pGRR-W8 of repression so the RGR-W20 could express more gRNA-W20 that inhibit the GFP expression. As a result, we can tune fluorescence back to OFF by adding the corresponding ligand. By this method, both Tet and Theo can be built into non-Boolean logical gate functioned as predicted.

The GFP fluorescence intensity is mainly controlled by the amount of W20 gRNA which is firstly driven by a constant pGDP promoter. It means the total amount W20 is steady and the increase of fluorescence is the amount of RGR-w20 that We can also predict the transfer function of this 2-layer RGA system by plotting the RGA activities against the RGR and compare with the experimental response functions.

3.2 Methods

Two-layer RGA Strain construction

Based on the yeast construction method in chapter 1, we substituted the pGRR-w8 promoter in reporter MATalpha strain with pGRR-w17w20 and integrated a pGRR-w8-RGR-w20 plasmid into it. The MATa strain stay the same as one-layer RGA and mate with each other with the same protocol

Testing two-layer RGAs

For two-layer cascades, three biological replicates have been tested with a selective Theophylline (0, 10, 20 mM) and Tetracycline (0, 0.4, 0.8, 1.2, 1.6 mM) series in eight β-estradiol series (0.01, 0.5, 1, 3, 8, 20, 100, 500 nM) according to the calibrating experiment and one-layer RGA testing experiments.

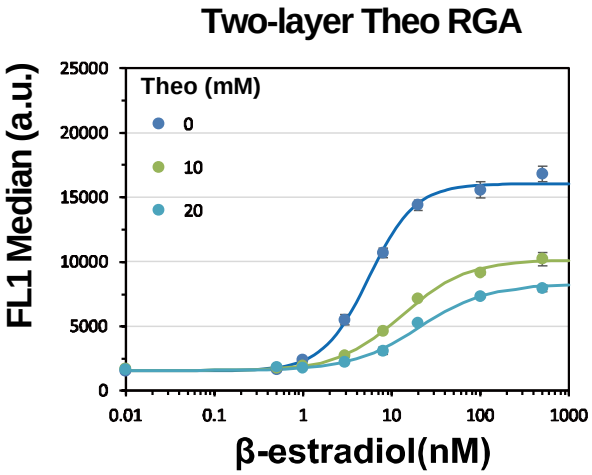
3.3 Experimental Results

Multi-layer cascade

We next tested two networks with more complex structures. Two-layer Theo and Tet can perform like an NIMPLY gate with RGA-W8 integrated and this two-layer structure

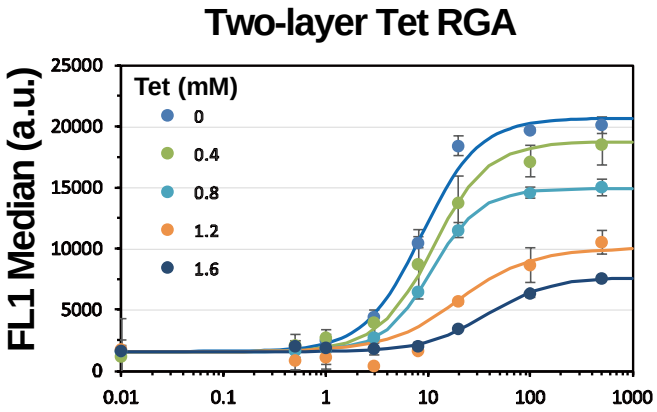
has an inverse behavior of the one-layer device(Figure 3.1-2). Only when the first input is high, the output can be high, otherwise, the signal will be low. Meanwhile, this structure gets a wide window that adding ligands can get 2.1X and 2.6X fold change in fluorescence when we compare the highest ligand concentration with the zero ligand conditions at high bE induction level. Here, we introduce the concept of inhibition ratio to evaluate the power of ligands to repress the corresponding aptazymes more precisely. In each bE induction, we assume the changes of the fluorescence between zero ligand and different titration conditions are all caused by the inhibition of aptazymes, so the ratio of this change to them maximum change available is the ratio of unfunctional RGAs inhibited by ligand to the total existing RGAs. Therefore, in each bE induction level, we can generate a transfer function of inhibition ration vs. ligand concentration which describes the ability of ligand to repress the aptazyme cleavage. We can also compare the sensitivity of different aptazymes or same aptazyme in different architectures.

In the two-layer device, the absolute fluorescence is larger than the two-layer Theo device which demonstrates tet RGA to have a faster cleavage rate than theo. We can also see a breakdown region around 1 mM. The fluorescence would be lower than the negative control after calibration may be due to the saturation of Tet in cells. And the fitting Hill function’s vMax value will be larger than 1 in some bE induction level, that also urges us to use Tet concentration lower than 1 mM.



	Ka	n	vMax	vMin
0	5.535223	1.668461	16062.12	1545
10	13.10791	1.22193	10139.74	1545
20	20.00328	1.139897	8247.296	1545

Fig 3.1 The Tetracycline and β-estradiol responsive dynamics plot of two-layer Tet-RGA cascade; the table consists of parameters from Hill equation fitting curves of RGA fluorescence data



	Ka	n	vMax	vMin
0	8.623468	1.550903	20686.94	1600
0.4	11.30905	1.608077	18764.39	1600
0.8	10.92696	1.817389	14901.14	1600
1.2	20.49077	1.222423	10072.75	1600
1.6	38.4444	1.429496	7642.048	1600

Fig 3.2 The Tetracycline and β-estradiol responsive dynamics plot of two-layer Tet-RGA cascade; the table consists of parameters from Hill equation fitting curves of RGA fluorescence data

Chapter 4: Multi-input RGA design

4.1 Circuit design

Since we have two different RGA as genetic input, we could now build multiple inputs systems with different architectures by arranging different RGAs in the same layer of network or in different layers. We still use pGal promoter to tune the RGA expression level and to analyze the prerequisite of building a Boolean or non-Boolean logic gates with different RGA inputs.

For devices whose both inputs in the same layer, like 2RGA1 and 2RGA2, we insert the RGA sequence into different selectable loci in the yeast genome and select the right haploid strain after double transformation. Then we mate this strain with two RGAs with a two-layer cascade pGRR17,20-yeGFP reporter with a pGRR8-RGR-w20 invertor. The signal of those systems should not be determined by only one ligand but by both, since the other one would still express effective gRNA when we only inhibit one RGA with one ligand.

While for device such as 2RGA12 whose inputs are in different layers, we designed it based on different responsive regime of one-layer and two-layer devices so we expect it to show response to different ligand at the same time. However, the overall expression level is hard to predict because the opposite effect of each RGA on the GFP reporter.

The advantage of genetic circuits over the digital ones is its nonlinear ligands binding behavior. We can modulate the speed of signal transition and which type of logical gate they give us by changing the ligand concentration matrix in the buffer.

4.2 Methods

2RGA1, 2RGA2, 2RGA12 yeast construction

Based on the yeast construction method in chapter 1, we insert the second RGA cassettes into different loci in the yeast genome with a different selective marker in reporter MATalpha strain. The MATa strain stay the same as one-layer RGA and two-layer RGA and we mate with each other respectively with the same protocol.

Testing two-inputs RGAs

For two-layer cascades, three biological replicates have been tested with four selective Theophylline and Tetracycline (0, 0.25, 0.5, 0.75 mM) series in eight β-estradiol series (0.01, 0.5, 1, 3, 8, 20, 100, 500 nM) according to the calibrating experiment and previous data. For 2RGA1 we used two Theo series (0, 4,8, 12 mM) and (0, 2, 4, 6 mM). For 2RGA1 we used two Theo series (0, 3, 6, 9 mM) and (0, 2, 4, 6 mM).

4.3 Experimental Results

Two RGA in the first layer (AND gate)

For 2RGA1, only when we add the mixture of both ligands will cause the fluorescence to increase dramatically, as the data showed. First, we titrated the device overall bE concentration and find the Ka and n is smaller than both of the single RGAs(Fig 4.1). It means the same bE induces two RGAs simultaneously. And the decrease of n value means

it exists little competence between those two kinds of gRNA-w8 expressed by different kinds of RGAs. The overall fluorescence is determined mainly by the RGA with the slowest cleavage rate and highest inhibition ratio while the other RGA contributes to the decrease of the total amount of gRNA, thus relieving the GFP from repression and increasing the fluorescence. It can also be reflected by the inhibition ratio that the gross inhibition ratio is always larger than the sum of two single inhibition ratios unless one of the ligands reaches saturation point or the addition of both inhibition ratio is close to one. Meanwhile, the single RGA inhibition ratio is always smaller than that of one-input RGAs which supports our assumptions. In low concentration of only +Tet/only +Theo condition, the fluorescence drops a little because the pGal promoter is leaky. When the one kind of free RGA is inhibited by a really small amount, the other one will take over the target without competence and repress GFP expression more easily and generates a more OFF state than the control. Despite the little repression caused by this effect, we can get great fluorescence activation by adding both ligands. Therefore, 2RGA1 can show a successful Boolean logical AND gate behavior that only both inputs are high the output value will be high. To solve the problem caused by the saturation of Theo inhibition, we use a smaller Theo concentration and same Tet concentration and it gives us a 2.5X fold increase of inhibition ratio(Fig 4.2).

Two RGA in the second layer (partial Boolean logic gate)

For 2RGA2, the overall Hill function is a combination of both single RNAs' activity. The Ka and n are between these parameters of the one-input RGAs' curves, which indicates the integration of two RGA is not a simple addition of both RGA activity but a complicated consequence of their cooperation and competition. Because of the instruction of an RGR layer between RGA and reporter, the GFP level is limited by the output of RGR rather than the output of RGA (Fig 4.3). Therefore, the signal of RGA will be delayed by RGR and cause the effect of competition larger than that of 2RGA1, which makes the signal RGA inhibition ratio much small than their original features. Just as we discussed in 2RGA1 device the total gRNA expressed will saturate earlier than any of the single RGAs. Moreover, because the saturation point will be close to the uninduced curves (Fig 4.4-5), so it will also be in the middle of the saturation point of single Tet and Theo RGAs. When we compare the inhibition ratio of 2RGA2 with the sum of inhibition ratio of two single RGAs, there must be a cross point between their curve plotted via inhibition ratio against ligand concentrations (Fig 4.6). Therefore, there are two different conditions due to the nonlinear response from our constructs.

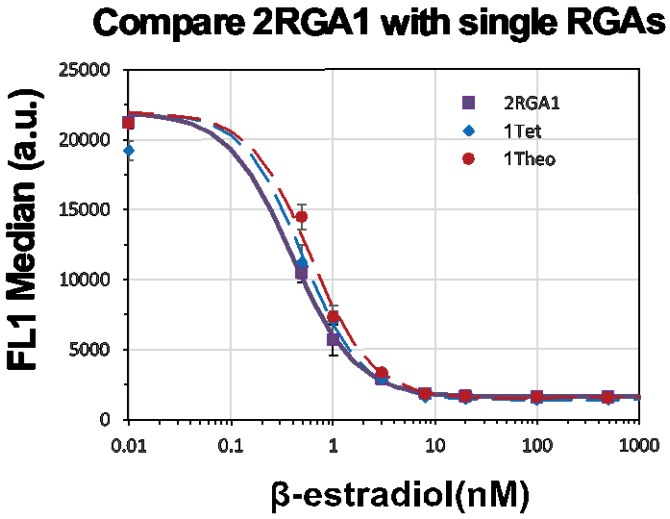
When we choose a relatively high Theo concentration, there is a large difference between Theo's inhibition ability than Tet's, so the overall inhibition ratio curve of 2RGA2 vs. ligand concentration will be close to Theo's transfer function. Also, it's easy to be saturated due to the use of large Theo value. As a result, the cross point exists after 2RGA2's saturation point and only when we use ligand concentration that smaller than the cross point, we can get the ideal response like a NAND Boolean logic gate (Fig 4.7). After the crossing point, the Boolean logic will break and the overall signal will always in the middle of single Theo and single Tet. In the other circumstance, when we choose a relatively low Theo concentration that has similar inhibition ability as Tet, the system won't be so easy to get saturate because Theo and Tet both contribute to the transfer function of 2RGA2

simultaneously. In this situation, the crossing point between transfer functions of 2RGA2 and the sum of two single RGAs is located before saturation point, so after the crossing point, 2RGA2 will always give us ideal NAND gate signal (Fig 4.8). In addition, 2RGA2 can give us imperfect Boolean logical gate before crossing point when the signal of 2RGA2 will be lower than two separate RGAs while is smaller than the sum of them.

Two RGAs in first and second layer separately (conditional Boolean logic gate)

Finally, we insert two RGAs with different gRNA target sequence into a different layer of our RGA network. The overall behavior is similar to the one-layer system that starts to form a high fluorescence ON state to an OFF state. In high bE level, all first layer RGA has been inhibited but there are still some RGR-W20 haven't been fully repressed by second layer RGA, then the fluorescence would increase a little like a two-layer device (Fig 4.9)

This system shows response to both ligands respectively without cooperation or competition. What's more, in different RGA expression level, the network will be able to sense different metabolites. In low bE induction level, the fluorescence will show an increase and restore full back to ON state and there will be little response to Tet. While in the intermediate range of induction, 2RGA12 can show activation with Theo and repression with Tet simultaneously and it displays a NIMPLY logic. When we apply a high level of bE induction, 2RGA12 won't be able to sense Theo anymore, but can still show response to Tet.



	Ka	n	vMax	vMin
2RGR1	0.393972	-1.39196	21973.5	1555
1Tet	0.507157	-1.50451	21973.5	1404
1Theo	0.602465	-1.47093	21973.5	1495

Fig 4.1 The β-estradiol responsive dynamics plot of One-layer Tet-RGA/Theo-RGA/2RGA1 cascade; the table consists of parameters from Hill equation fitting curves of RGA fluorescence data.

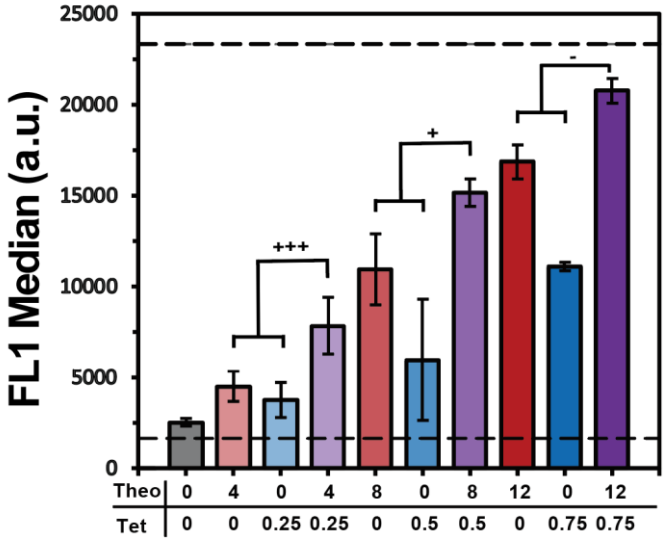
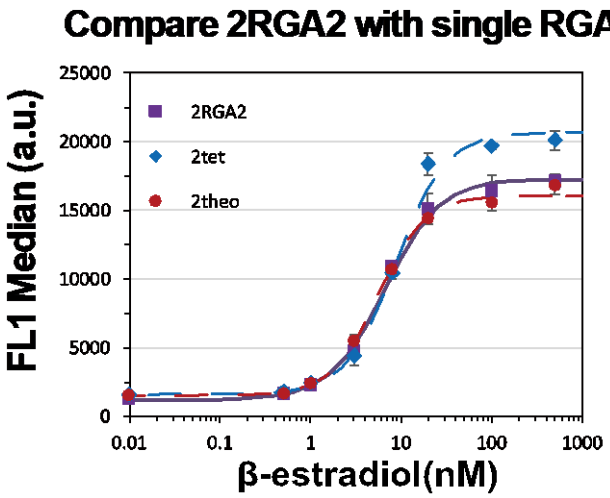


Fig 4.2 The fluorescence response plot of 2RGA1 device in mixture buffer with selective Theo and Tet concentration at 1 nM bE induction level, compare the inhibition ratio of 2RGA1 with the addition of each component.



	Ka	n	vMax	vMin
2RGR2	6.6706	1.43627	17302.75048	1232.5
2Tet	8.623468	1.550903	20686.94	1600
2Theo	5.535223	1.668461	16062.12	1545

Fig 4.3 The β-estradiol responsive dynamics plot of two-layer Tet-RGA/Theo-RGA/2RGA2 cascade; the table consists of parameters from Hill equation fitting curves of RGA fluorescence data.

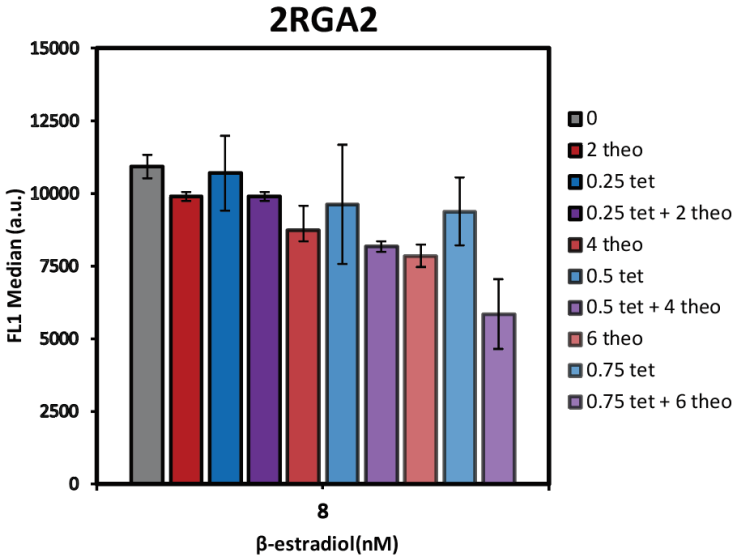


Fig 4.4 The fluorescence response plot of 2RGA2 device in mixture buffer with selective Theo and Tet concentration at 8 nM bE induction level.

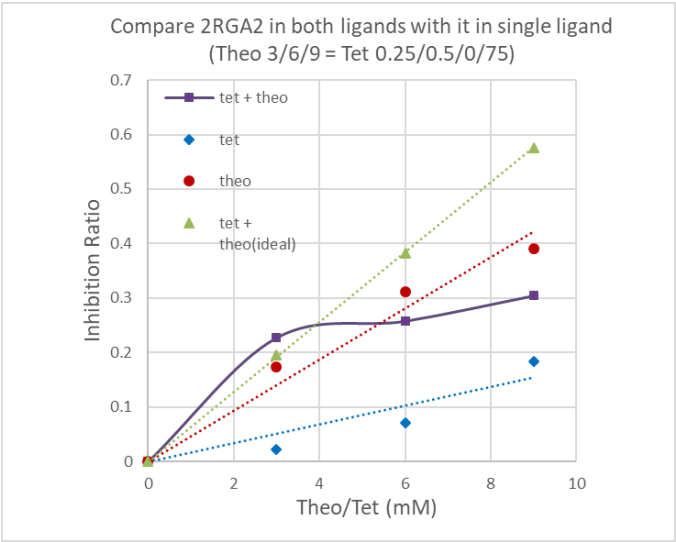


Fig 4.5 The inhibition ratio plot of 2RGA2 device in mixture buffer with selective Theo (0, 3, 6, 9 mM) and Tet (0, 0.25, 0.5, 0.75 mM) concentration at 8 nM bE induction level, compare the inhibition ratio of 2RGA1 with the addition of each component.

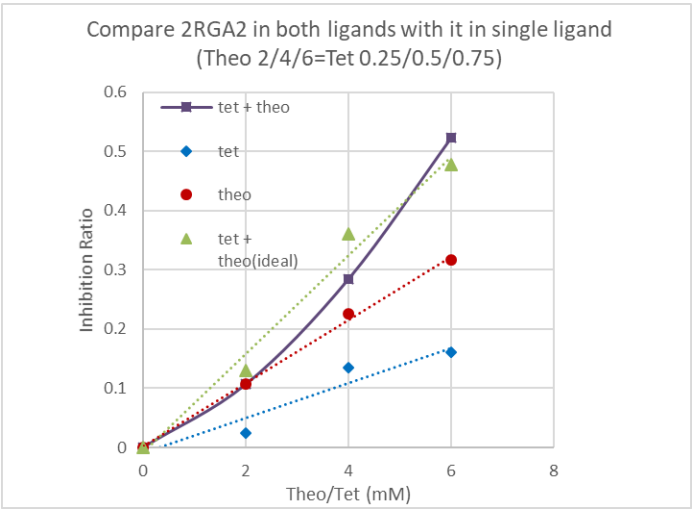


Fig 4.6 The inhibition ratio plot of 2RGA2 device in mixture buffer with selective Theo (0, 3, 6, 9 mM)and Tet (0, 0.25, 0.5, 0.75 mM) concentration at 8 nM bE induction level, compare the inhibition ratio of 2RGA1 with the addition of each component.

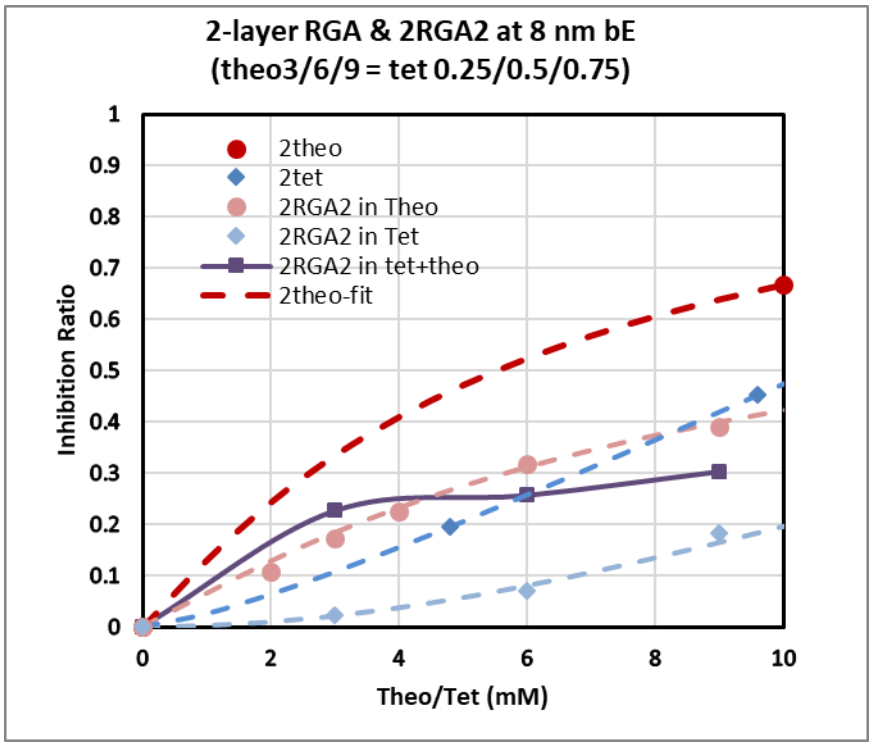


Fig 4.7 The inhibition ratio plot of 2RGA2 device in mixture buffer with selective Theo (0, 3, 6, 9 mM)and Tet (0, 0.25, 0.5, 0.75 mM) concentration at 8 nM bE induction level, compare the inhibition ratio of 2RGA1 with with two-layer RGAs and each component in 2RGA2.

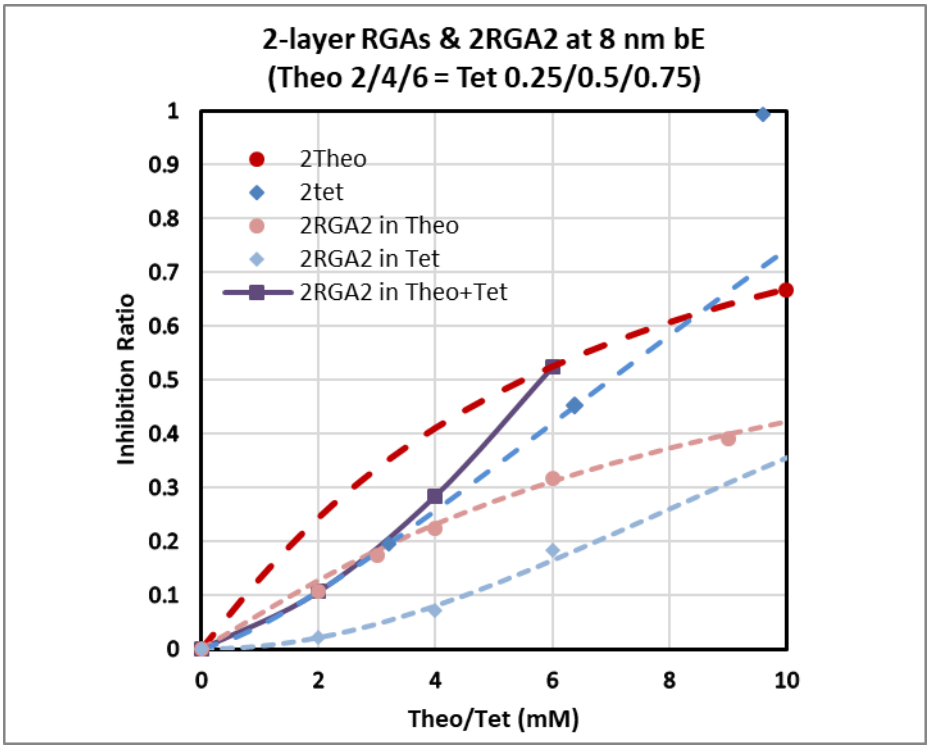


Fig 4.8 The inhibition ratio plot of 2RGA2 device in mixture buffer with selective Theo (0, 2, 4, 6 mM)and Tet (0, 0.25, 0.5, 0.75 mM) concentration at 8 nM bE induction level, compare the inhibition ratio of 2RGA1 with two-layer RGAs and each component in 2RGA2.

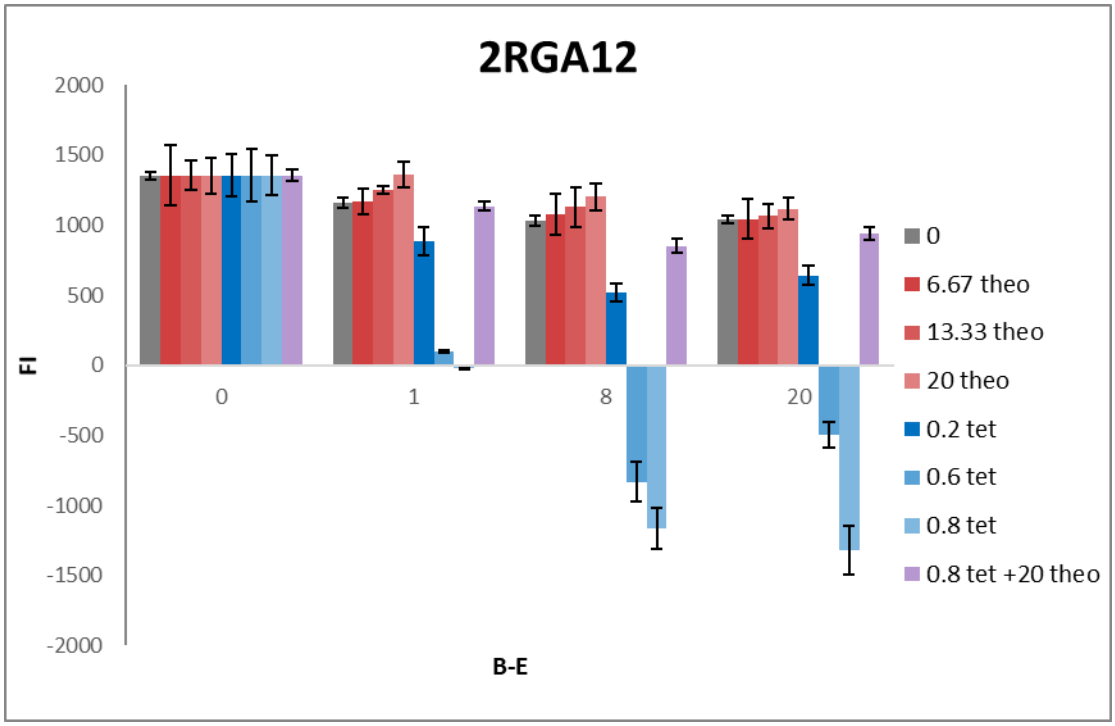


Fig 4.9 The fluorescence response plot of 2RGA1/2 device in mixture buffer with selective Theo and Tet concentration at 1, 8 and 20 nM bE induction levels. Use 0 bE induction condition as a positive control.

Chapter 5: Conclusions

In this work, we constructed three types of multi-Ribozyme-gRNA-Aptazyme (RGA) networks for biosensing and gene regulation by link RGA, Ribozyme-gRNA-Ribozyme (RGR) and CRISPRi systems together. We screened out two different aptazymes that fit this network and analyzed their one-layer and two-layer single-input RGAs and functions in two-inputs systems with different architectures. We demonstrate the RGA device can work as what we designed. With the introduce of different aptamers, we can compensate the influence of β-estradiol induction by tuning the ligands concentration. Most importantly, our first two-layer RGA network has been characterized so this series of architectures promises that we can design more complicated circuit like what we've done with RGRs to engineer metabolic pathways.

In multi-input systems, can still get repression signal if we add relative high concentration of single ligand, but only when both inputs are at the similar repression level by corresponding ligand, the overall fluorescence can be changed dramatically. The requirement to realize a successful Boolean logic gate form multi-RGA system is 1) comparable inhibition ratio value when single RGA gets saturated 2) use the concentration before this RGA system saturate. We analyzed the responsive patterns of each multi-input networks and give explanations of their characteristics to make them more predictable.

Overall, we built two IMPLY gates, two NIMPLY gates, one AND gate, one partial NAND gate, one conditional non-Boolean gate. The RGA shows as much as a 7.38-fold increase in ligand-dependent system-output. The interaction between RGAs that are inserted in circuits in parallel has also been reported, which contributions to the predictability of more complicated RGA networks.

REFERENCES

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science (80-.). 337, 816–821.

Nissim, L., Perli, S.D., Fridkin, A., Perez-Pinera, P., and Lu, T.K. (2014). Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells. Mol. Cell 54, 698–710.

Weber, J., Valiante, V., Nødvig, C.S., Mattern, D.J., Slotkowski, R.A., Mortensen, U.H., and Brakhage, A.A. (2017). Functional reconstitution of a fungal natural product gene cluster by advanced genome editing. ACS Synth. Biol. 6, 62–68.

Fei, J., Knapp, D., Schuez, M., Murawala, P., Zou, Y., Pal Singh, S., Drechsel, D., and Tanaka, E.M. (2016). Tissue- and time-directed electroporation of CAS9 protein–gRNA complexes in vivo yields efficient multigene knockout for studying gene function in regeneration. Npj Regen. Med. 1, 16002.

Gander, M.W., Vrana, J.D., Voje, W.E., Carothers, J.M., and Klavins, E. (2017). Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. Nat. Commun. 8, 15459.

Yoshioka, S., Fujii, W., Ogawa, T., Sugiura, K., and Naito, K. (2016). Development of a mono-promoter-driven CRISPR/Cas9 system in mammalian cells. Sci. Rep. 5, 18341.

Briner, A.E., Donohoue, P.D., Gomaa, A.A., Selle, K., Slorach, E.M., Nye, C.H., Haurwitz, R.E., Beisel, C.L., May, A.P., and Barrangou, R. (2014). Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality. Mol. Cell 56, 333–339.

Mekler, V., Minakhin, L., Semenova, E., Kuznedelov, K., and Severinov, K. (2016). Kinetics of the CRISPR-Cas9 effector complex assembly and the role of 3′-terminal segment of guide RNA. Nucleic Acids Res. 44, 2837–2845.

Chuai, G. hui, Wang, Q.L., and Liu, Q. (2017). In Silico Meets In Vivo: Towards Computational CRISPR-Based sgRNA Design. Trends Biotechnol. 35, 12–21.

Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2015). Rationally engineered Cas9 nucleases with improved specificity. Science (80-.). 351, 84–88.

Wong, N., Liu, W., and Wang, X. (2015). WU-CRISPR: Characteristics of functional guide RNAs for the CRISPR/Cas9 system. Genome Biol. 16, 1–8.

Liu, Y., Zhan, Y., Chen, Z., He, A., Li, J., Wu, H., Liu, L., Zhuang, C., Lin, J., Guo, X., et al. (2016a). Directing cellular information flow via CRISPR signal conductors. Nat. Methods.

Tang, W., Hu, J.H., and Liu, D.R. (2017). Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. Nat. Commun. 8, 1–8.