RICE UNIVERSITY

Transcriptional delay in synthetic genetic cascades

by

Yu-Yu Cheng

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Matthew R. Bennett
Associate Professor, Department of Biosciences

Krešimir Josić
Associate Professor, Department of Mathematics
University of Houston

George N. Phillips
Professor, Department of Biosciences

Jeffrey J. Tabor
Assistant Professor, Department of Bioengineering

Weiwei Zhong
Assistant Professor, Department of Biosciences

HOUSTON, TEXAS
April 2017

ABSTRACT

Transcriptional delay in synthetic genetic cascades

by

Yu-Yu Cheng

Transcription factors (TFs) and their target promoters are central to synthetic biology. By arranging these components into complex regulatory networks, synthetic biologists have been able to create a wide variety of phenotypes, including bistable switches, oscillators, and logic gates. However, transcription factors do not instantaneously regulate downstream targets. After the gene encoding a TF is turned on, it must first be transcribed, the transcripts must be translated, and sufficient TF must accumulate in order to bind operator sites of the target promoter. The time to complete this process, here called the "transcriptional delay," is a critical aspect in the design of dynamic regulatory networks, yet it remains poorly characterized. In this work, I measured the delay of two TFs in *Escherichia coli*, which are commonly used in synthetic biology: the activator AraC and the repressor LacI. I found that the delay can range from a few to tens of minutes, and are affected by the expression rate of the TF. The single-cell data also shows that the variability of the delay increases with its mean. To validate these time measurements, I constructed a two-step genetic cascade, and showed that the timing of the full cascade can be predicted from those of its constituent steps. These results demonstrate the timescale of transcriptional regulation in living cells, which is important for understanding the dynamics of synthetic transcriptional gene

circuits.

Acknowledgments

I would like to thank Matt for giving me the opportunity and freedom to explore this field of research. In addition, I am grateful for the tremendous support from the lab. This work would not have been possible without the help of many people. I would like to specially thank Andrew Hirning, David Shis, Ye Chen, Alan Veliz-Cuba, and Krešimir Josić.

The love from my family got me here. This work is dedicated to them.

Contents

List of Figures

Nomenclature

aa	Amino acid
ARA	L-arabinose
aTc	Anhydrotetracycline
BCD	Bicistronic design
bp	Base pair
CRISPRi	Clustered regularly interspaced short palindromic repeats interference
E. coli	*Escherichia coli*
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny broth
mRNA	Messenger RNA
nt	Nucleotide
OD	Optical density
ODE	Ordinary differential equation
PCR	Polymerase chain reaction
PDMS	Poly-dimethylsiloxane
RBS	Ribosome binding site
RNase	Ribonuclease
sfYFP	Super-folder yellow fluorescent protein
sgRNA	Short-guide RNA
shRNA	Short-hairpin RNA

SLIC	Sequence and ligation independent cloning
TF	Transcription factor

Chapter 1

Introduction

1.1. Synthetic genetic circuits

Two papers published back-to-back in the *Nature* in January 2000 marked the start of the era of synthetic biology. One paper demonstrated the oscillatory behavior of a negative-feedback system composed of three transcriptional repressors inhibiting the production of one another (Elowitz and Leibler, 2000). The other constructed a memory system with two repressors, each repressor controlling the expression of the other (Gardner et al., 2000). The idea that synthetic genetic circuits can be engineered like electronics drew a lot of attention. The human genome was sequenced and published the next year (Consortium, 2001; Venter et al., 2001). The focus was transitioning from "reading DNA" to "writing DNA." More circuits were built since then, such as a pulse generator (Basu et al., 2004), an event counter (Friedland et al., 2009), logic gates (Moon et al., 2012), analog calculators (Daniel et al., 2013), and an edge detector (Tabor et al., 2009). The idea to engineer cells for different applications took hold. Some examples are

using engineered cells for producing chemicals from raw materials (Nielsen and Keasling, 2016), killing cancer cells (Din et al., 2016; Xie et al., 2011), and detecting soil contamination (Prindle et al., 2012). With the increasing knowledge of circuit construction, one day synthetic genetic circuits may be applied *in vivo*.

However, before achieving precise organismal engineering, there are many problems that need to be solved. First, unlike electronic components, biological parts are far more heterogeneous. Currently, there is no general rule to assemble parts. It usually takes 3–5 years to build a functional circuit. Hence, studying the rules of part assembly is important. Second, the available parts are limited. With more parts available, we can design complex circuits to do sophisticated tasks. Third, the speed of circuit construction is slow. Researchers in laboratories often order short DNAs from commercial companies to clone the target circuits from existing parts. Cloning techniques can assemble DNA together and avoid synthesizing large DNA strands. However, creating a prototype often takes weeks or even months. Without clear rules to assemble parts, researchers need to experiment using a lot of trial-and-error. Next, I will discuss possible solutions towards solving these problems.

Concerning difficulty in part assembly, controlling the translation rate of proteins can be used as an example. The translation rate at the ribosome binding site (RBS) is often “context-dependent.” The secondary structure of RNA of the upstream promoter and downstream protein coding region will affect the rate of ribosome binding (Espah Borujeni et al., 2014). Thermodynamics models have been proposed to calibrate the translation rate, but the accuracy is still low (Li, 2015). To solve this problem, a standard RBS design has been implemented to accommodate different promoters and proteins

(Mutalik et al., 2013). With two RBS in parallel, the interference from upstream sequence is minimal. There are also reports on the standardization of promoter (Brewster et al., 2012) and transcription terminator (Chen et al., 2013b). The expression level of proteins is especially important for transcription factors (TFs). TFs regulate the transcription of promoters by either recruiting or blocking RNA polymerase. By controlling TF's expression from other TFs, these logic elements can be used to construct functional modules. However, the production rate of TF must be accurately controlled in a certain range. The goal of my work is to measure the time delay for TFs to regulate the promoter in this concentration window. I will address the importance of the signaling time of TFs in the next section.

The second problem is the scarcity of available parts. Parts can be sorted into three categories: input, processor, and output. The input elements are sensors for the information outside or inside of the cell. The information can be chemical such as sugars (Shis et al., 2014), a gas molecule (Prindle et al., 2012), or a basic chemical element (Prindle et al., 2012). It can also be physical, such as temperature (Hoynes-O'Connor et al., 2015) and light (Levskaya et al., 2005). The processor is composed of molecules which are able to execute logic operation, such as TFs or specifically designed RNAs (Chappell et al., 2015a). The output layer components are usually proteins to deliver the specific functions. To expand the repertoire of these genetic parts, we can develop bioinformatics tools to explore the genome of different organisms. For example, through DNA sequence comparison, dozens of TFs were mined from *E. coli*-related bacteria (Stanton et al., 2014). We can also generate random sequences and select useful ones. 265 synthetic transcription terminators have been reported from a large randomly

generated library (Chen et al., 2013b). Another way is to design molecules with computational models. For example, a synthetic protein has been designed to inhibit the infection of influenza virus (Fleishman et al., 2011).

To accelerate the speed of circuit construction, one possible solution is to improve DNA synthesis technology. It is now possible to order 2000 bps of DNA from one commercial company (Integrated DNA Technologies, Inc.), the DNA is still shorter than the average size of a typical plasmid. Once this limit is overcome, testing circuit designs can be much faster, and the tedious labor involved in this process can be avoided. Another way is to build DNA libraries. There are now facilities dedicated to the storage and sharing of DNA, such as the Registry of Standard Biological Parts and Addgene. When the exchange of parts becomes easier, circuit designs can be tested quickly by simply incorporating existing tools.

In summary, efforts are being made toward advancing synthetic genetic circuit design. With the feasibility of synthesis of long DNA fragments, standard part assembly methods, and powerful bioinformatics tools, we will use engineered cells to cure currently incurable diseases and generate valuable materials from microorganisms in the future.

1.2. Delay in synthetic genetic circuits

Most synthetic genetic circuits use TFs in the circuit design. However, the time delay associated with regulation of the promoter by TFs is currently an unknown parameter. In this study, I aimed to measure this time delay. The delay is important for

genetic circuits with dynamical properties, such as logic gates and oscillators. For example, the delay of logic gates must be comparable to avoid faulty outputs (Moon et al., 2012). Also, it is thought that delay is important for generating robust oscillation for negative-feedback circuits (Mather et al., 2009; Stricker et al., 2008). In bistable systems, longer delays may make the steady state more stable (Gupta et al., 2013). In addition to the time scale of the delay, I also aimed to measure the distribution of the delay among isogenic cells. Since it has been shown that gene expression in single cells is heterogenous, it is possible that the delay is distributed rather than being a constant. The measured distribution could help refine models for genetic circuits. For example, distributed delay could affect the propagation of information in gene cascades (Josić et al., 2011). Finally, I also aimed to study how delay could affect expression of downstream genes. It is possible that gene expression and delay are correlated due to the extrinsic noise (Elowitz et al., 2002).

Transcriptional regulation involves many steps, from the transcription and translation of the TF gene, protein folding and oligomerization to the final steps of promoter searching and binding (Figure 1).

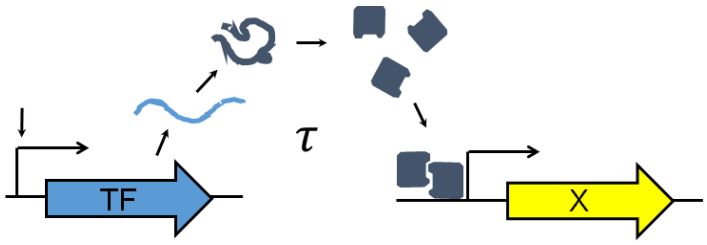

Figure 1. Transcriptional regulation involves many steps. The delay is the summation of reaction times of transcription, translation, protein folding, protein oligomerization, and TF binding of DNA.

All of the steps are stochastic and are affected by dilution, degradation, and various sources of cellular noise (Pedraza and van Oudenaarden, 2005). The timescale of some steps has been measured, such as the transcription rate of RNA polymerase (Vogel and Jensen, 1994), translation rate of ribosomes (Bremer and Dennis, 1987), and the promoter-searching rate of TFs along DNA (Elf et al., 2007). In the next section, I will illustrate these reactions and their measured rates. However, the overall timescale and the variability of the delay are not clear, which I aimed to measure in this study. To measure the delay of transcription regulation, I built an activation circuit and a repression circuit to measure delay for both transcriptional activation and repression. A two-step cascade was also built to examine how the delay is accumulated in multi-step transcriptional regulations.

Several synthetic genetic cascades have been built (Dunlop et al., 2008; Hooshangi et al., 2005; Olson et al., 2014; Pedraza and van Oudenaarden, 2005; Rosenfeld and Alon, 2003). The cascade built by Rosenfeld and Alon showed a 50 minute delay for transcriptional repression (Rosenfeld and Alon, 2003). The cascade built by Hooshangi et al. showed a delay over 100 minutes in a two-step genetic cascade (Hooshangi et al., 2005). The cascade built by Dunlop et al. showed a 120 minute delay in only transcriptional repression (Dunlop et al., 2008). Olson et al. used an optogenetics approach to measure the time delay for the repressor TetR to repress P_{Tet} promoter, and they showed that the time is 7.0 ± 5.4 minutes (Gardner et al., 2000). The cascade built by Pedraza and van Oudenaarden showed that the stochasticity from upstream TFs can be propagated to downstream gene expression, but did not provide dynamic information (Pedraza and van Oudenaarden, 2005). It seems that there is no consensus about the delay

of transcriptional regulation. In this study, I used a microfluidic device to trigger the built circuits and observe single cell fluorescence changes, which is more precise than the systems used in these mentioned studies.

Genetic cascades are a common motif in naturally occurring transcriptional networks (Rosenfeld and Alon, 2003). For simple organisms, such as *E. coli*, the typical length of a gene cascade is one or two steps. In contrast, higher organisms tend to have longer cascades. For example, the genetic cascades in *Drosophila* can be as long as nine steps. It is possible that single cell organisms need fast response times to cope with environmental changes, thus limiting the development of long genetic cascades. For multicellular organisms, longer cascades can be used to coordinate developmental processes. Measurement of the delay will reveal the precise timing of genetic cascades. For example, Stavens et al. found that the variation of the delay decreased along the lytic pathway of bacteriophage λ and hypothesized that more precise timing could be achieved by a longer cascade (Amir et al., 2007).

1.3. Steps of transcriptional regulation

The steps of transcriptional regulation are illustrated below. The measured reaction rate is emphasized here to give a view about how long the overall reaction will take.

Inducer import

In this study, I used IPTG and ARA to induce gene expression. IPTG is an analog of allolactose, which triggers the transcription of the *lac* operon. IPTG is not metabolized

by beta-galactosidase from the *lac* operon and can induce gene expression from the P_{lac} promoter. IPTG is transported into *E. coli* cells by the lactose permease from the *lac* operon, and other pathways are also involved (Marbach and Bettenbrock, 2012). At low IPTG concentrations, IPTG can enter cells only through lactose permease. Also, when IPTG concentration is lower than a certain threshold (~30 μM), gene expression of the *lac* operon appears to be bimodal (Ozbudak et al., 2004). It is suggested that the bimodality arises from the positive feedback of gene regulation in the *lac* operon (Choi et al., 2008).

The metabolism of the sugar ARA is via the *araBAD* operon. ARA enters cells through the transporters AraFGH and AraE from the *araBAD* operon (Megerle et al., 2008). Like IPTG, when ARA concentration is lower than a certain threshold (~1.33 mM or 0.02%), cells will appear to be bimodal (Siegele and Hu, 1997). The timing of P_{BAD} gene expression has been studied under different ARA concentrations (Megerle et al., 2008). It has been shown that gene expression can be detected as soon as four minutes. When the ARA concentration is reduced, the timing of gene expression increases. The variability of timing also increases. However, it is difficult to tell whether the increase in timing is a result of the uptake of ARA or the slow accumulation of reporter proteins.

Gene synthesis and degradation

Gene synthesis begins with the initiation of transcription. When the inducer triggers the conformational change of the TF, the TF will either recruit or block the binding of RNA polymerase to the promoter and start the transcription. It is estimated that there are 1,000–10,000 RNA polymerase per cell (Bremer and Dennis, 1987). The

speed of transcription is about 40–80 nt/sec (Bremer and Dennis, 1987; Proshkin et al., 2010; Vogel and Jensen, 1994). The ribosome will then recognize the RBS on the mRNA and start the translation. It is estimated that there are 10,000–100,000 ribosomes per cell (Bremer and Dennis, 1987). The speed of translation is about 20 aa/sec (Bilgin et al., 1992; Bremer and Dennis, 1987; Proshkin et al., 2010), which corresponds to 60 nt/sec. For a 1 kb gene in *E. coli*, the completion of transcription should take 25 seconds at most. Translation will take roughly the same time. Thus, gene synthesis can be completed within 1 min after the initiation of transcription.

The concentration of RNAs and proteins will be continuously diluted by cellular growth. If cells divide every τ_{cyc} minutes, the concentration will be halved after τ_{cyc} minutes. Assume that the molecules are rather stable, the concentration can be described as follows:

$$\frac{\mathrm{dC}}{\mathrm{dt}} = -\mathrm{k} * \mathrm{C}$$

where C is the concentration of the molecule and the rate constant $\mathrm{k} = \frac{\ln(2)}{\tau_{cyc}}$. However, the RNAs and proteins are not simply diluted; they are also actively degraded by enzymes. It is estimated that RNA has half-life times between 3–8 minutes (Bernstein et al., 2002). Proteins are more stable. Even when tagged with a degradation motif, the half-lifetime of proteins is still roughly the same as the cell doubling time (Andersen et al., 1998).

Protein maturation

The linear polypeptide chain starts to fold into its three-dimensional structure once it has been synthesized. The first step is the formation of secondary structures, such as an alpha helix or a beta sheet, which is driven by intramolecular hydrogen bonds. Then, the three-dimensional shape of the protein is formed, which is called the tertiary structure. Formation of the tertiary structure is driven by hydrophobic interactions between the secondary structures and the aqueous environment. Once the tertiary structure is formed, formation of disulfide bonds between cysteine residues strengthen the structure. Finally, multiple folded proteins could aggregate together to form functional modules, which is termed the quaternary structure. In this study, the activator AraC is a dimer (Bustos and Schleif, 1993), and the repressor LacI is a tetramer (Swint-Kruse and Matthews, 2009).

The protein folding timescale is often much shorter than one second (Gromiha et al., 2006). Less is known about the timescale of protein aggregation, for which concentration is important. For proteins that need further reactions to mature, the time will be longer. For example, the chromophore inside fluorescent proteins needs to undergo an oxidation reaction to become functional. The measured maturation time of fluorescent proteins can range from less than five minutes to an hour, depending on organisms, temperature, and the reporter construct (Gordon et al., 2007; Nagai et al., 2002). The maturation time is especially important for the measurement of response time for biological processes.

Transcription factor dynamics

Once a transcription factor is formed, it starts the search for a specific DNA sequence. It first binds nonspecifically to a DNA strand and then moves randomly to search the target. However, this 1D searching will take a long time to find the target. A TF will only scan about 85 bp, then it will jump to other strands to start a new search. Theoretically the combination of 1D and 3D diffusions allow TF to search the target more efficiently (Li et al., 2009). It was estimated that it will take one TF molecule at most six minutes to find an operator site (Elf et al., 2007). If there are more than one TF molecules in the cell, the search time can be reduced to one minute, which is the diffusion-limited search time. Once the TF binds to the target sequence, it will recruit or block the RNA polymerase.

Based on these measured reaction rates, the time delay for TF to regulate the promoter can be as fast as a few minutes, which mostly come from the delays of transcription, translation, and DNA search. At lower ligand concentrations, delay could increase due to longer ligand uptake time and DNA search time. For synthetic genetic circuits, the TF and the target promoter are often built on plasmids. Depending on the replication origin, the copy number of plasmids can range from one to hundreds. When the number of promoter is more than one, we need to consider the time that most of promoters are bound with TF molecules, which will depend on both the production rate and degradation rate of TF. Also, different TFs have different binding affinity to the promoter. If the binding affinity is low, the TF may need more time to reach an effective concentration. In this study, I measured the delay for two commonly used TFs – AraC and LacI, and characterized how the delay increases when lower inducer concentration is used.

1.4. Mathematical modeling and stochastic simulation

To model synthetic genetic circuits, ordinary differential equations (ODEs) are often used to describe the concentration change of molecules inside the cell. The rate of concentration change is determined by the production and degradation terms. The production terms are usually determined by the synthesis processes, such as transcription and translation. The degradation terms are based on the dilution from cellular growth and enzymatic degradation (Andersen et al., 1998). Once the equations and parameters are set, one can find the steady-state solutions of the system, study how the initial condition affects the final state, and determine how the parameters could affect the outcome (Strogatz, 1994). ODE models are simple enough to describe the behavior of genetic oscillators (Elowitz and Leibler, 2000; Stricker et al., 2008) and the toggle switch (Gardner et al., 2000), two of first circuits built. The ODEs can also be computed using software such as MATLAB if the model is too complex to solve analytically.

Sometimes computational modeling can guide circuit design. In the process of modeling a robust oscillator (Stricker et al., 2008), it was found that it is only possible to describe the oscillatory behavior of a dual-feedback oscillator by incorporating each detailed reaction. Hence, a small delay from sub-steps is required to generate the robust oscillations (Stricker et al., 2008). This hypothesis was confirmed by building a negative-feedback only version of the circuit. Modeling also helps the synchronization of single-cell oscillators (Nathan et al., 2016) and tuning of the period of oscillators (Din et al., 2016).

However, the behavior of circuits at the single-cell level has often been found to be heterogeneous. This stochasticity can be attributed to the small copy number of reactants such as DNA (intrinsic noise) or the heterogeneous states of the cell (extrinsic noise). These two factors have been shown to play equally important roles in gene expression (Elowitz et al., 2002). To take the randomness of the system into account, chemical master equations or Langevin equations are often used to describe circuits' behavior (van Kampen, 2007). The stochastic model has been used to study the noise propagation in genetic cascades (Paulsson, 2004; Pedraza and van Oudenaarden, 2005) and the stochastic behavior of *Synechococcus elongatus* circadian clocks (Chabot et al., 2007). The stochastic model can be simulated by algorithms such as Gillespie algorithm (Gillespie, 1976; Gillesple, 1977) or tau-leaping method (Rathinam et al., 2003).

The goal of this research is to investigate the importance of delays in modeling synthetic genetic circuits. Delays have been shown to nontrivially change the dynamical behavior of a negative-feedback circuit (Mather et al., 2009; Stricker et al., 2008). A delay differential equation shows that a small delay could lead to oscillation with a much longer period (Mather et al., 2009). A general form of delay differential equations with a constant delay can be written as follows:

$$\frac{d}{dt}x(t) = f(t, x(t), x_\tau),$$

where $x_\tau = x(t-\tau), \tau > 0$ is the trajectory of the solution in the previous time interval.

If the delay is distributed rather than constant, the equation can be modified as follows:

$$\dot{\mathrm{x}}(t) = \int_0^{\mathrm{t}} \mathrm{g}[\mathrm{x}(t), \mathrm{x}(t-s)]\mathrm{f}_{\mathrm{T}}(s)\mathrm{d}s,$$

where $\mathrm{f}_{\mathrm{T}}(s)$ is the probability of the distributed delay. In addition to differential equations, queuing models from stochastic process study can also be applied. For example, a queuing model has been used to explain the correlated gene expression when two genes are degraded by a limited number of ClpXP enzymes (Cookson et al., 2011). Recently it has also been used to study the signaling time of genetic circuits (Josić et al., 2011). Numerical methods can also incorporate either a constant delay or a distributed delay (Barrio et al., 2006; Bratsun et al., 2005; Xiaodong, 2007). This study aimed to measure the distribution of delay among isogenic cells and determine whether new delay models are needed for synthetic genetic circuits.

Chapter 2

Research Methods

2.1. Genetic circuit construction

2.1.1. Synthetic genetic circuits and strains

In this study, six plasmids and three *E. coli* stains were used to build seven synthetic genetic circuits. The details of the seven synthetic genetic circuits are described in the following section.

Activation circuit

The TF gene AraC was placed under the $P_{A1lacO-1}$ promoter (Lutz and Bujard, 1997) (Figure 2). AraC was tagged with an LAA sequence (Andersen et al., 1998). The RBS for AraC is B0034 (http://parts.igem.org/Part:BBa_B0034). The super-folder gene YFP (sfYFP (Kremers et al., 2006)) was placed under the P_{BAD} promoter. The RBS for

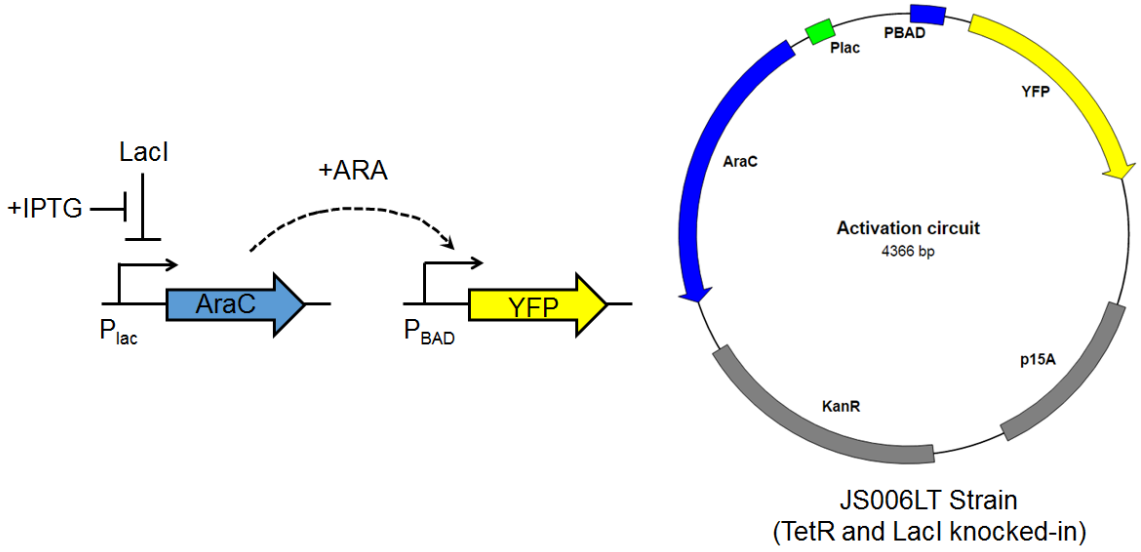

Figure 2. The activation circuit. The activation circuit was constructed by placing the AraC gene under control of the P_{lac} promoter and the YFP gene under control of the P_{BAD} promoter. The plasmid was transformed into JS006LT cells, which constitutively express LacI to repress the P_{lac} promoter. When IPTG and ARA are added, AraC will activate YFP.

yfp is a bicistronic design (BCD) (Mutalik et al., 2013). The activation plasmid has a kanamycin resistant gene and a p15A origin. The plasmid was transformed into JS006LT cells, which harbor TetR and LacI constitutively expressed from strong promoters. All strains used were derived from the JS006 strain (Stricker et al., 2008). The circuit was triggered with 2 mM, 0.2 mM, or 0.05 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) separately, plus 2% L-arabinose (ARA).

Repression circuit

The TF gene LacI was placed under P_{BAD} promoter (Figure 3). LacI was tagged with an LAA sequence. The RBS for LacI is BCD. The super-folder YFP gene was

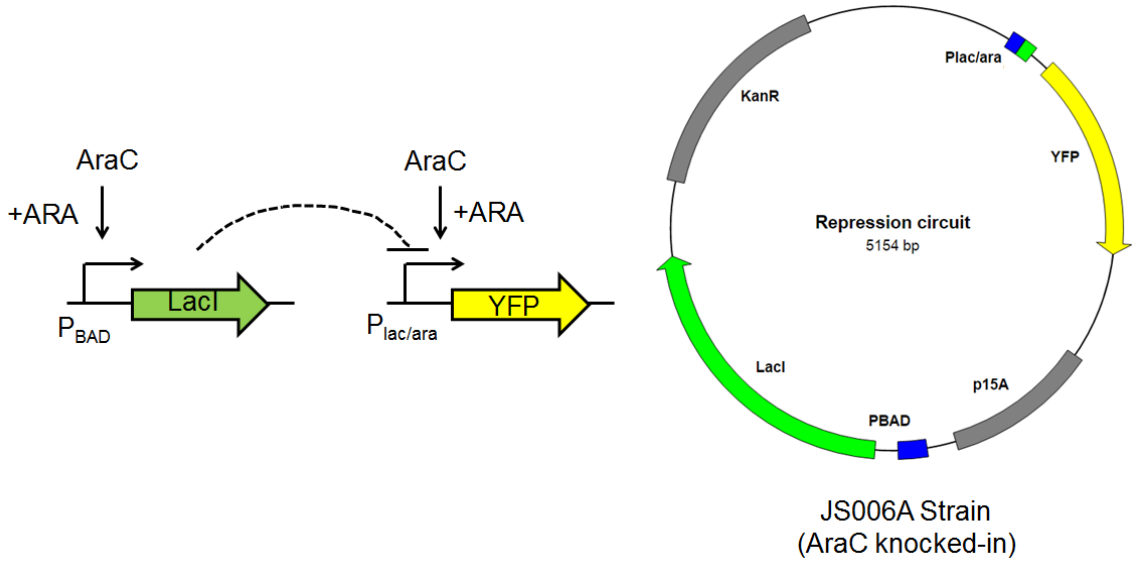

Figure 3. The repression circuit. The repression circuit was constructed by placing the gene LacI under control of the P_{BAD} promoter and the YFP gene under control of the $P_{lac/ara}$ promoter. The plasmid was transformed into JS006A cells, which constitutively expresses AraC to activate P_{BAD} and $P_{lac/ara}$ promoters. When ARA is added, AraC will activate both LacI and YFP. Then, LacI will repress YFP.

placed under control of the $P_{lac/ara}$ promoter (Lutz and Bujard, 1997). YFP was tagged with an LAA sequence. The RBS for *yfp* is BCD. The repression plasmid has a kanamycin resistant gene and a p15A origin. The plasmid was transformed into JS006A cells, which have AraC constitutively expressed from a strong promoter. The circuit was triggered with 2%, 0.1%, 0.05%, and 0.02% ARA separately.

Two-step genetic cascade

The circuit is composed of two plasmids: the P_{Tet}-AraC plasmid and the repression plasmid (Figure 4). To construct the P_{Tet}-AraC plasmid, the TF gene TF AraC was placed under control of the P_{Tet} promoter (Lutz and Bujard, 1997). The RBS for

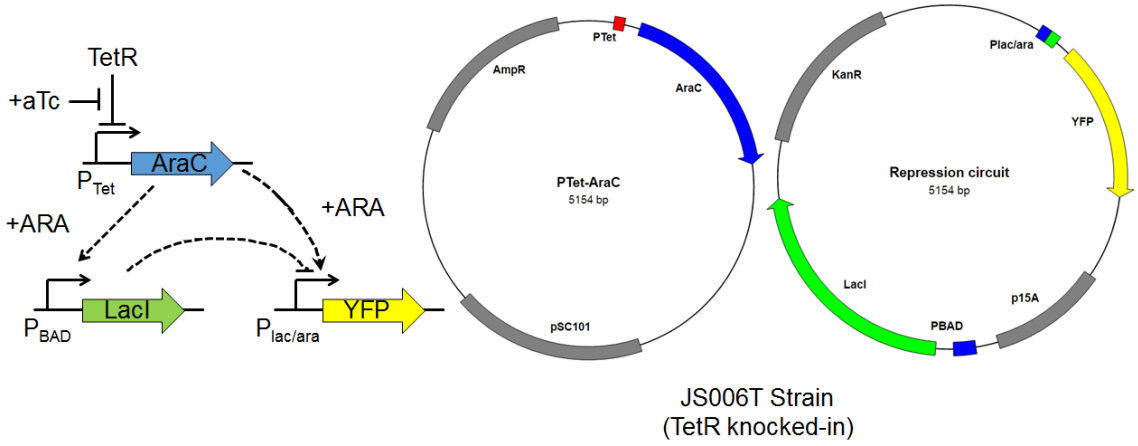

Figure 4. The two-step genetic cascade. The cascade comprises the activation and repression steps. The gene AraC was placed under control of the P_{Tet} promoter, the LacI gene under control of the P_{BAD} promoter, and the YFP gene under control of the $P_{lac/ara}$ promoter. The two plasmids were transformed into JS006T cells, which constitutively express TetR to repress the P_{Tet} promoter. When aTc and ARA are added, AraC will activate both LacI and YFP. Then, LacI will repress YFP.

AraC is BCD. AraC was tagged with an LAA sequence. The P_{Tet}-AraC plasmid has an ampicillin resistant gene and a pSC101 origin. The two plasmids were co-transformed into JS006T cells, which have TetR constitutively expressed from a strong promoter. The circuit was triggered with 1 μg/ml anhydrotetracycline (aTc) and 2% ARA.

Modified activation circuit

The circuit is composed of two plasmids: the P_{Tet}-AraC plasmid and the P_{BAD} reporter-only plasmid (Figure 5). To construct the P_{BAD} reporter-only plasmid, the sfYFP gene was placed under control of the P_{BAD} promoter. RBS for YFP is BCD. The P_{BAD} reporter-only plasmid has a kanamycin resistant gene and a p15A origin. The two

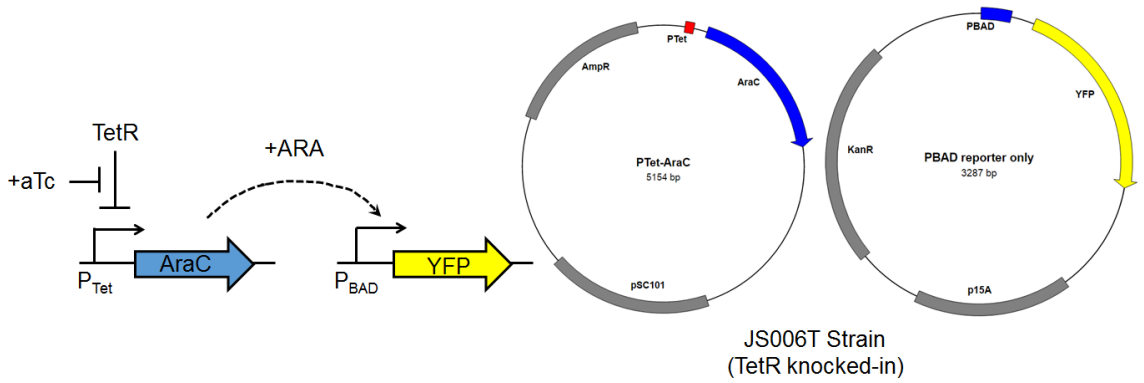

Figure 5. The modified activation circuit. The modified activation circuit was constructed by placing the gene AraC under control of the P_{Tet} promoter and the YFP gene under control of the P_{BAD} promoter. The two plasmid were transformed into JS006T cells, which constitutively express TetR to repress the P_{Tet} promoter. When aTc and ARA are added, AraC will activate YFP.

plasmids were co-transformed into JS006T cells. The circuit was triggered with 1 μg/ml anhydrotetracycline (aTc) and 2% ARA.

P_{BAD} reporter-only circuit

The P_{BAD} reporter-only plasmid was transformed into JS006A cells (Figure 6A). The circuit was triggered with 2% ARA. The circuit was also used for the characterization of P_{BAD} promoter activity under different ARA concentrations.

P_{lac} reporter-only circuit

The sfYFP gene was placed under control of the P_{lac} promoter (Figure 6B). YFP was tagged with an LAA sequence. The RBS for *yfp* is BCD. The P_{lac} reporter-only

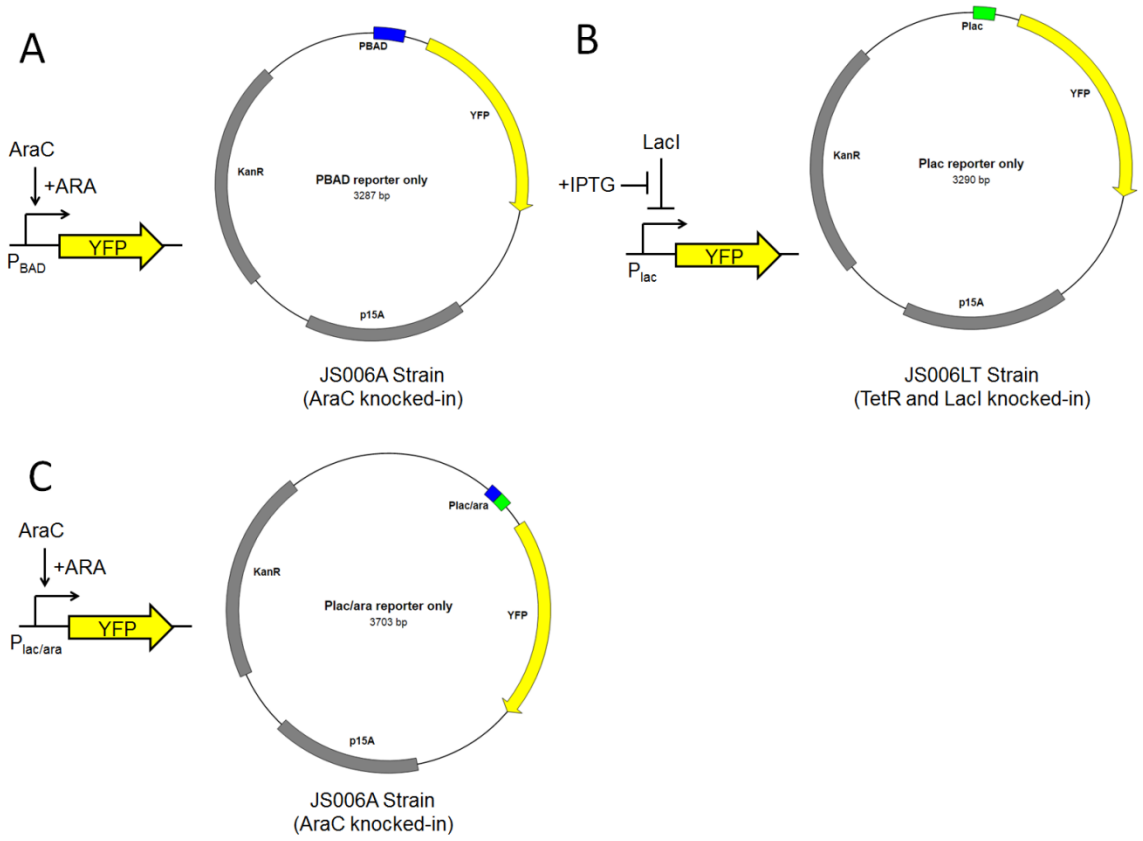

Figure 6. The reporter-only circuits. (A) The P_{BAD} reporter-only circuit was constructed by placing the YFP gene under control of the P_{BAD} promoter. The plasmid was transformed into JS006A cells, which constitutively express AraC to activate P_{BAD}. When ARA is added, AraC will activate YFP. (B) The P_{lac} reporter-only circuit was constructed by placing the YFP gene under control of the P_{lac} promoter. The plasmid was transformed into JS006LT cells, which constitutively express LacI to repress P_{lac}. When IPTG is added, YFP will be expressed. (C) The $P_{lac/ara}$ reporter-only circuit was constructed by placing the YFP gene under control of the $P_{lac/ara}$ promoter. The plasmid was transformed into JS006A cells, which constitutively express AraC to activate $P_{lac/ara}$. When ARA is added, AraC will activate YFP.

plasmid has a kanamycin resistant gene and a p15A origin. The plasmid was transformed into JS006LT cells. The circuit was used for the characterization of P_{lac} promoter activity under different IPTG concentrations.

$P_{lac/ara}$ reporter-only circuit

The sfYFP gene was placed under control of the $P_{lac/ara}$ promoter (Figure 6C). YFP was tagged with an LAA sequence. The RBS for *yfp* is BCD. The $P_{lac/ara}$ reporter-only plasmid has a kanamycin resistant gene and a p15A origin. The plasmid was transformed into JS006A cells. The circuit was used for the characterization of $P_{lac/ara}$ promoter activity under different ARA concentrations.

2.1.2. Cloning and gene knock-in methods

Two cloning methods to construct the plasmids were used: PCR mutagenesis and Sequence and Ligation Independent Cloning (SLIC) (Li and Elledge, 2012). PCR mutagenesis allows for short sequence changes in the plasmids, such as the promoter, the LAA tag, and the RBS. SLIC was used to combine two DNA fragments and to insert genes into plasmids.

PCR mutagenesis uses two DNA primers to change sequences in the plasmid. The primers have two parts: an annealing part and a new part (Figure 7A). The annealing part is used to amplify the plasmid from the region of interest. The new part has the modified sequence that overlaps with the other primer. The annealing parts and the overlapping part should have melting temperature (T_m) around 70°C. The T_m was calculated using the NetPrimer website (http://www.premierbiosoft.com). During the PCR, the primers will amplify the plasmid with the new part. The PCR product can be directly transformed into competent cells for sequence confirmation. The strain of competent cells used in this work was DH10BALT, which have RecA knocked out and have AraC, LacI, and TetR knocked in.

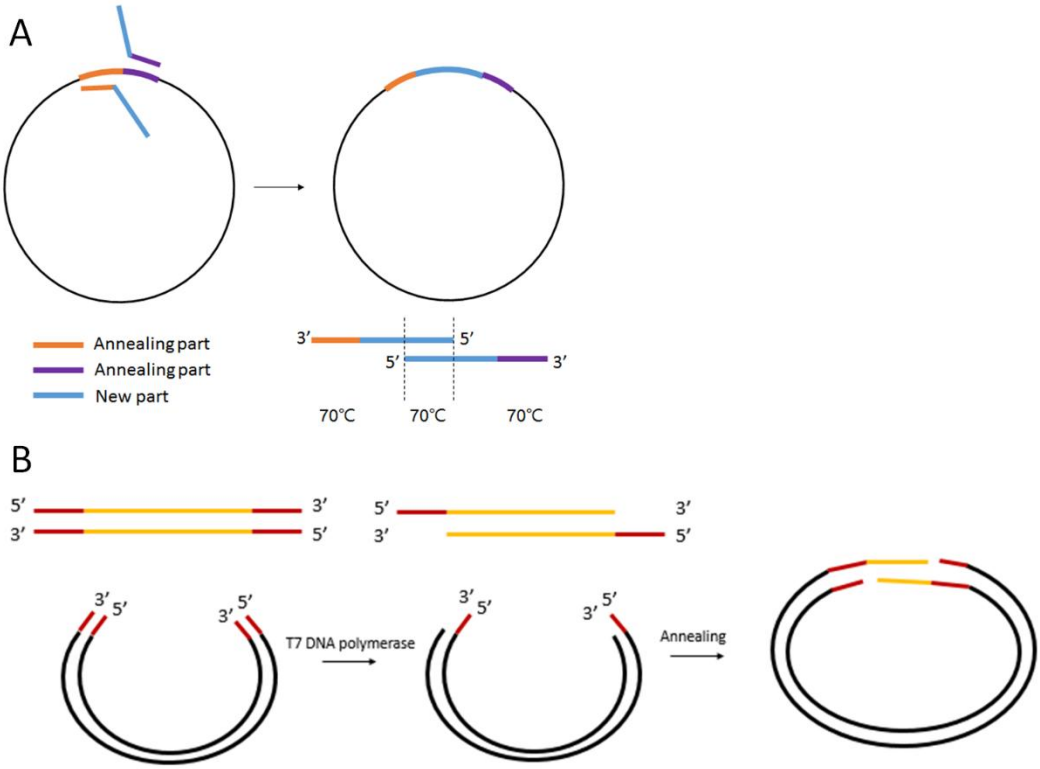

Figure 7. Cloning methods. (A) PCR mutagenesis. To modify plasmids, two primers were designed to introduce the changes. The primers contain an annealing part and a new part. The annealing part is used to amplify the plasmid and has a sequence homologous to the plasmid. The new part is the modified sequence, which overlaps with the other primer. The T_m should be approximately 70°C. (B) Sequence and ligation independent cloning. To join two linear DNA fragments, T7 DNA polymerase is used to remove nucleotides at the 3′ end of the double-stranded DNA. The DNA fragments then will anneal to each other through the introduced 20 bp homologous sequence.

SLIC can be used to join two DNA fragments together (Figure 7B). The two fragments are PCR amplified with 20 bp at the ends. This 20 bp homology can be introduced by primers. After gel purification, the linear DNA fragments were treated with T7 DNA polymerase, which excises the 3' end of DNA fragments. The reaction is then stopped by adding dGTP. Next, the two fragments are mixed together to allow them to anneal to each other. The final product is then ready for transformation.

The three *E. coli* strains used in this study, JS006LT, JS006A, and JS006T, were derived from the JS006 strain, which has *araC* and *lacI* knocked out (Stricker et al., 2008). The genes were knocked in using bacteriophage lambda integrase (Diederich et al., 1992). The inserted genes were first cloned into a vector with the lambda attachment site *attP*. Next, the origin of the vector was cut out and the linear DNA was re-ligated to form circular DNA. The target strain was transformed with a helper plasmid, which contains the integrase. The expression of the integrase is repressed by a temperature-sensitive cI gene. The circular DNA was then transformed into the target strain with the heat shock method. Growing the cells at 42°C caused the integrase to insert the genes into the *attB* site of the *E. coli* genome. The integration was checked by colony PCR.

2.1.3. Promoter characterization

To characterize the promoter activity, the fluorescence expression of three circuits, P_{BAD} reporter-only, P_{lac} reporter-only, and $P_{lac/ara}$ reporter-only, were measured with various inducer concentrations. Cells were cultured overnight and then diluted 1:100 and 200 μL was transferred to a 96-well plate. The inducer was added at this time. The cells were grown in a 37°C shaker for two hours to reach an OD of approximately 0.1–0.3. The YFP fluorescence and OD600 were measured using an Infinite® 200 PRO fluorescence reader. The fluorescence was divided by the OD600 value to generate the induction curve (Figure 8). Each data point is an average of three measurements. The $P_{lac/ara}$ promoter is approximately 10-fold stronger than the P_{BAD} promoter. The inducer

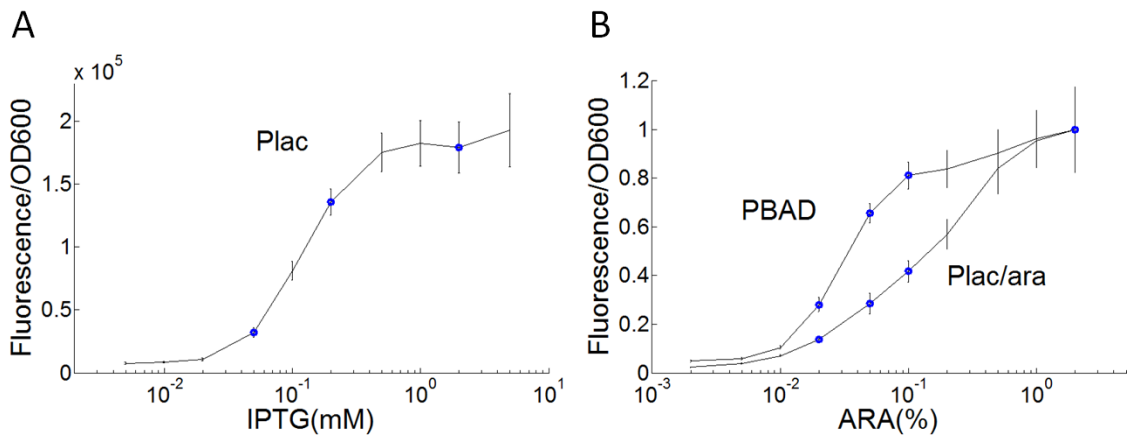

Figure 8. The induction curves of promoters P_{lac}, P_{BAD}, and $P_{lac/ara}$. (A) P_{lac} activity under different IPTG concentrations. Three IPTG concentrations were chosen to trigger the activation circuit. (B) P_{BAD} and $P_{lac/ara}$ activities under different ARA concentrations. Four ARA concentrations were chosen to trigger the repression circuit. Notice that P_{BAD} and $P_{lac/ara}$ respond differently to ARA concentration. P_{BAD} is approximately 10-fold stronger than $P_{lac/ara}$ (data not shown).

concentrations used to trigger the other circuits were based on the measured induction curves.

2.2. Microfluidic device and microscope experiment

2.2.1. Dial-a-wave chip

To measure the distribution of delays among isogenic cells, a microfluidic device was used to trigger circuits and to observe single cells under the fluorescence microscope. The 'dial-a-wave' chip was designed by Andrew Hirning for the purpose of this study (Figure 9). The chip is composed of five ports linked to syringes with different media. The media port was loaded with LB. The inducer port was loaded with LB and inducers as well as the dye Sulforhodamine 101 (Sigma-Aldrich). The two waste ports were

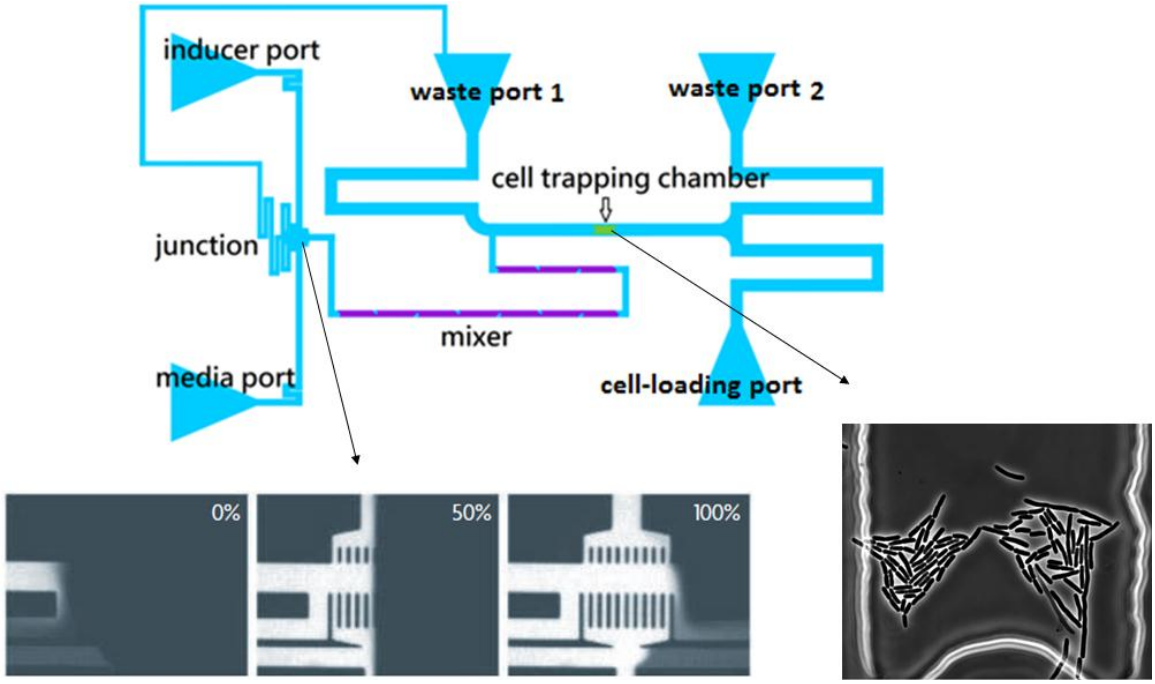

Figure 9. The "dial-a-wave" microfluidic chip. The chip is composed of five ports for loading media. The junction is where the inducer is calibrated. The mixer is a feature used to fine-tune the media composition. The cell trapping chamber is where single cells are grown and imaged. The calibration images were adapted from (Bennett and Hasty, 2009).

loaded with LB. The cell-loading port was used to insert cells into the device. The junction and mixer were used to fine-tune the media composition ranging from 0% inducer to 100% inducer. At the junction, the dye was used to calibrate the heights of the syringes feeding the media port and inducer port to produce ranges between 0% to 100% inducer. The cells were loaded into the cell trapping chamber, which is 100 μm wide, 300 μm long, and 0.95 μm high. The channels are 10 μm high. Cells in the chamber were subsequently photographed.

To make a PDMS chip from a wafer, first the PDMS base and catalyst were mixed at a ratio of 10:1. The wafer was cleaned with isopropanol and wrapped in

aluminum foil. Then the mix was poured onto the wafer. The gas inside the uncured PDMS was removed by vacuum pump before curing the chip at 80°C in an oven. After the PDMS was cured, the five ports were opened with a tissue puncher. Then, the PDMS was cut into individual chips and cleaned with methanol in an ultrasonic cleaner. The chips were baked at 80°C for another 30 minutes. After the coverslip was cleaned with ozone, the chip was then ready for binding with the coverslip. The final PDMS chips were baked overnight to enhance the binding.

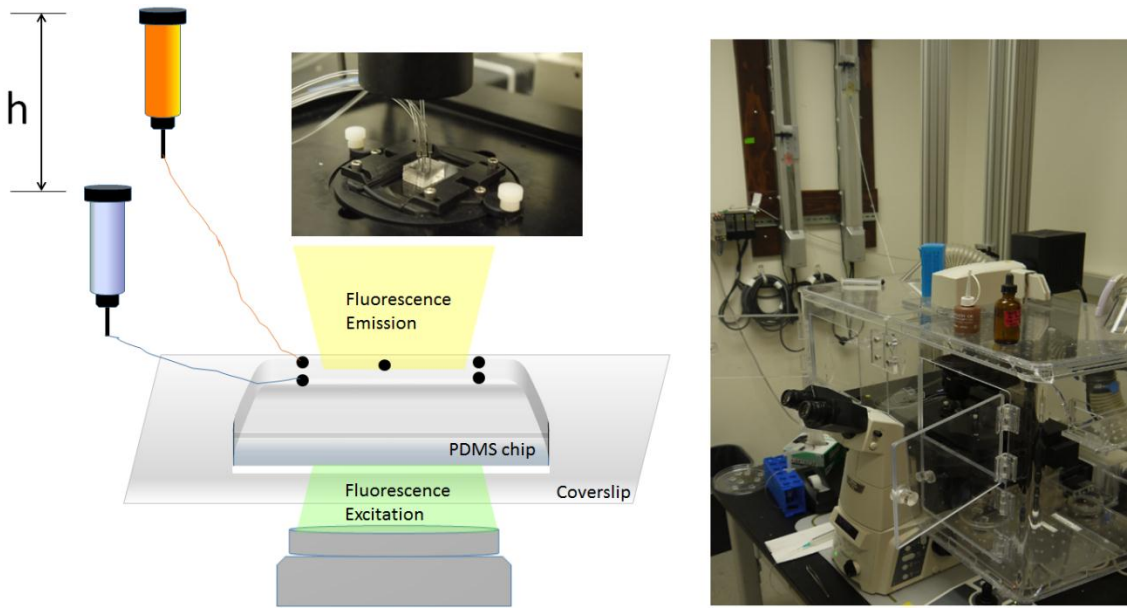

Figure 10. Microscope experiment. The chip was loaded with media from syringes, which were hung on columns. The heights of the syringes can be automatically adjusted and the difference in heights (h) will determine the composition of media inside the chip. The chip is mounted onto the stage of the fluorescence microscope, which take phase-contrast and fluorescence images. The temperature is kept at 37°C by an enclosed case and a heater.

2.2.2. Microscope experiment

Before conducting the microscope experiment, microfluidic device needed to be set up. First, the chip was mounted on the stage of the Nikon Eclipse Ti-E microscope. Then the chip was flushed with 0.1% tween, which prevents cells from adhering to the walls. It is important to make sure the cell trapping chamber is flushed. Then, all syringes can be loaded with media. For the waste port, 5 mL LB with 5 µL antibiotics were loaded into 20 mL syringes. For the media port, 10 mL LB with 10 µL antibiotics were loaded into 60 mL syringes. For the inducer port, 10 mL LB, 10 µL antibiotics, inducers, and 20 µL of 1 mg/mL dye were loaded into the 60 mL syringe. After the chip was mounted on the stage of the Nikon Eclipse Ti-E microscope, the syringes were connected to the chip via tubes and pins (Figure 10). The height of the waste port was kept low at all times. The heights of the media port and the inducer port needed to be calibrated using the dye (Figure 9). Then, the chip was ready for loading the cells.

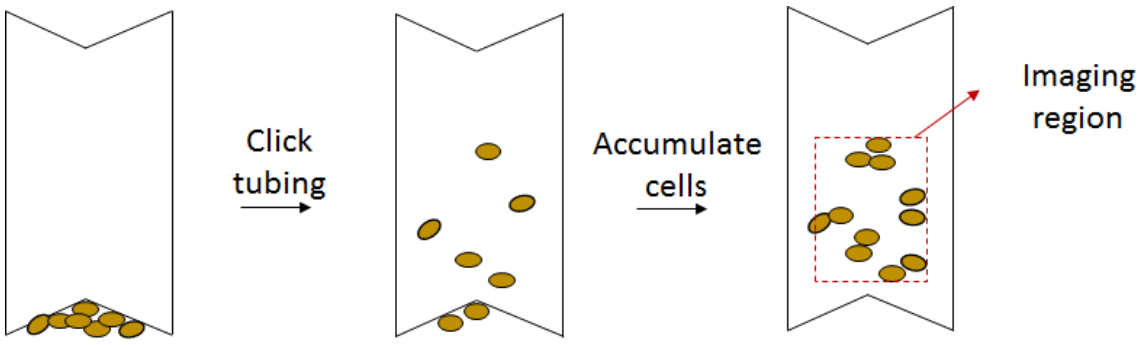

Figure 11. Loading cells into the cell trapping chamber. When cells are loaded into the chip, some of the cells get stuck at the edge of the chamber. To push the cells into the chamber, the operator can click the tubing. An instantaneous increase of the flow rate will push the cells into the chamber. Then, the operator waits for cells to accumulate to around ~40–50 cells to start the imaging.

After the characterization, the overnight culture of cells was diluted 1:50 into 5 mL LB for one hour to allow cells to resume the exponential phase. The cells were then loaded into the chip via the cell loading port. However, some cells would get stuck on the edge of the cell trapping chamber (Figure 11). Clicking the tubing can create extra pressure to push cells into the chamber. After the number of cells reached approximately 40–50, the imaging was started. Phase contrast (100× magnification with additional 1.5× magnification) and fluorescence images (mCherry and YFP channels) were taken every minute. A 20-watt excitation light was used for the fluorescence excitation. The exposure time for the YFP and mCherry were 300 ms and 100 ms, respectively. The inducer was switched to 100% after 10 minutes. The YFP signal in this 10-minute period was used as the background level to define the threshold for the activation of the circuit.

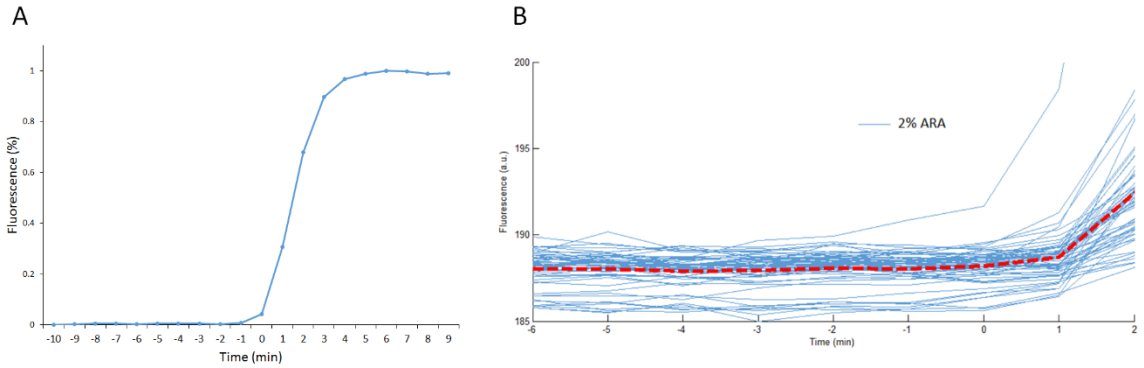

Figure 12. Definition of time for microscope experiments. (A) The dye trajectory for a step function input. It takes 5 min for the fluorescence level to reach maximum. This is limited to the mixing of the media in the chamber. (B) The time at which the first point of fluorescence increase was seen was set as time 0. For strong promoters, such as $P_{lac/ara}$, the fluorescence increase can be detected in 1–2 minutes, using this time reference.

Each experiment was performed twice to get at least 60 single-cell fluorescence traces for statistical analysis. The imaging was conducted over a two-hour period, and the images were used to determine single-cell YFP expression.

The dye signal from the mCherry channel was used to track the concentration of the inducers. Typical dye trajectories show that it takes approximately five minutes for the inducer to reach its maximal concentration (Figure 12A). The first time point that inducer started to increase was defined as time 0 (5%). I found that for YFP expressed from strong promoters, such as $P_{lac/ara}$ promoter, the YFP signal can be detected within 1–2 minutes (Figure 12B). The same time reference was used for all experiments.

2.3. Data analysis

2.3.1. Cell segmentation and fluorescence readout

To extract single cell fluorescence signals, the cells in the phase-contrast images needed to be segmented manually (Figure 13). mtPaint was used to facilitate the task by clicking two points to create a line segment (http://mtpaint.sourceforge.net/). When the cells divided, I randomly picked one sibling cell to segment. Then, the segmented images were analyzed by a semi-automatic tracking algorithm developed by Alan Veliz-Cuba. The algorithm can identify the same cells on the time-series of images according to their limited movement and using a one-to-one correspondence. It also allows for manual correction. After all the cells are identified, it reads out the fluorescence for each cell from the fluorescence images. The fluorescence value is the total fluorescence signal in the segmented region divided by the number of pixels. Therefore, it is proportional to the

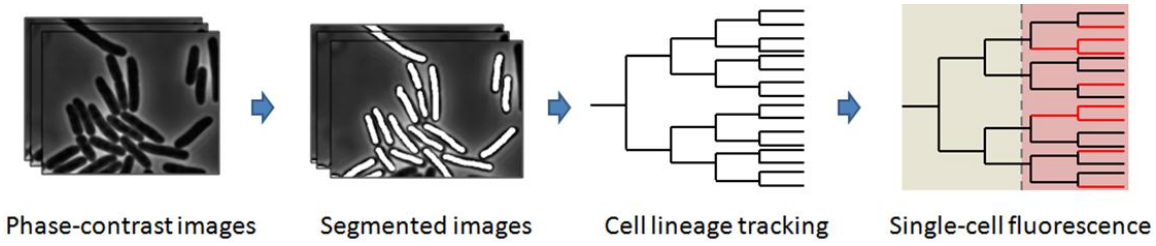

Figure 13. Image analysis. The cells in the phase-contrast images were first segmented manually. Then, a script tracked each cell's lineage along the time course. After the cell lineage was confirmed, the single-cell fluorescence was calculated from the fluorescence images.

YFP concentration inside the cell. The single-cell fluorescence data is used for further analysis. The script can be downloaded online (github.com/alanavc/rodtracker).

2.3.2. Calculation of P-values for the correlation test

In this work, I measured the delay of transcriptional regulation and the downstream gene expression of single cells. The correlation coefficients of each inducer concentration tested show that the two variables are negatively correlated, except for the short delay data. To test whether the two variables have a linear relationship, the null and alternative hypotheses are set as:

Null hypothesis H_0: $\rho = 0$

Alternative hypothesis H_A: $\rho \neq 0$

The test statistics for the t-test for testing the population correlation coefficient is $t = \frac{\rho\sqrt{n-2}}{\sqrt{1-\rho^2}}$, where n is the sample size. Then, the two-tailed P-value can be determined when t and n are known. It can be calculated using an online tool (https://graphpad.com/quickcalcs/pValue1/). A P-value is the probability of how likely it is that a given result will be obtained when the null hypothesis is true. If the P-value is small, it is likely that the two variables are correlated.

Chapter 3

Transcriptional delay in synthetic genetic cascades

In this section, the data from the built circuits will be presented. The activation circuit and repression circuit were used to measure the delay for transcriptional activation and repression, respectively. Mathematical models were also proposed to explain the measured delays. Finally, the two-step genetic cascade was used to examine how the delay is convolved when two circuits are combined together. The precision of the measurements was also tested.

3.1. AraC activation time

The activation circuit was used to measure the transcriptional activation of the P_{BAD} promoter by AraC (Figure 14A). The AraC gene was placed under the inducible promoter P_{lac}. YFP was placed under the control of the P_{BAD} promoter as a reporter.

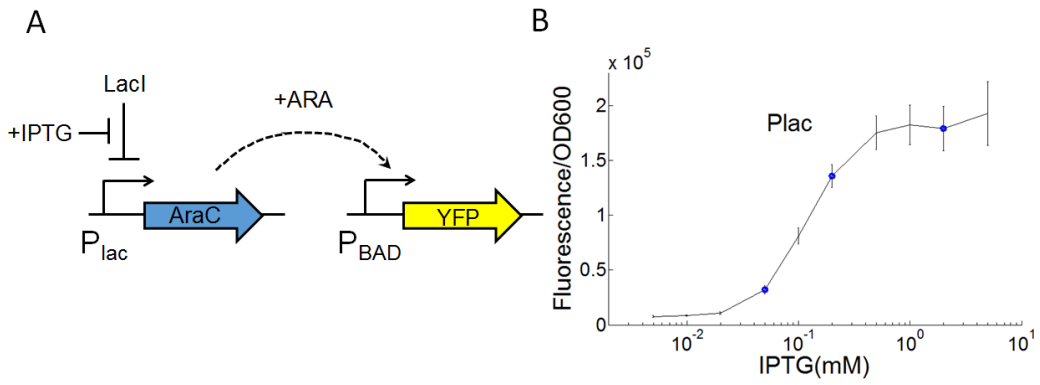

Figure 14. Measurement of activation delay. (A) The activation circuit. Adding IPTG will trigger the production of AraC. Then AraC will bind to ARA to turn on the YFP. (B) Three IPTG concentrations were tested on the circuit to see how the production rate of AraC would affect the activation delay.

Once IPTG and ARA are added, AraC will be expressed from the P_{lac} promoter, bind with ARA, and activate the P_{BAD} promoter. Then, YFP is expressed and the fluorescence signal inside cells increases. The circuit was tested with three different IPTG concentrations (Figure 14B). The P_{BAD}-only reporter circuit was also tested as a control (Figure 6). The single-cell data shows that when IPTG concentration is high, the YFP expression of the activation circuit is similar to the P_{BAD}-only reporter circuit (Figure 15). Reducing the IPTG concentration resulted in lower YFP expression, and YFP signals appeared later. At 0.05mM IPTG concentration, most cells did not express much YFP, but a few cells had high YFP expression.

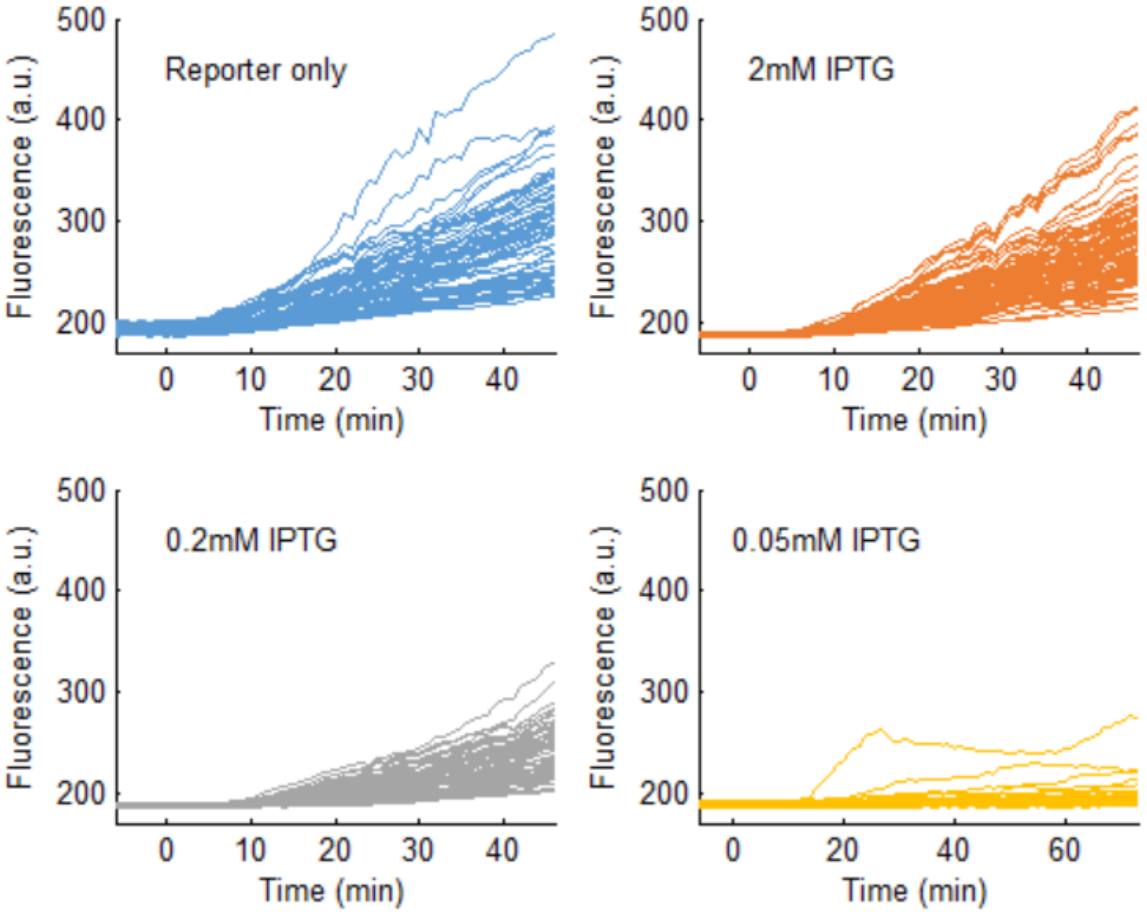

Figure 15. Single-cell fluorescence of activation circuit with 2 mM, 0.2 mM, and 0.05 mM IPTG plus 2% ARA, and P_{BAD} reporter-only circuit with 2% ARA. Each experiment was repeated twice and at least 60 cells were collected.

To estimate the delay of transcriptional activation, the background fluorescence signal was used to define a threshold for the activation time (Figure 16). The threshold was chosen to be the mean plus N standard deviations of the 10-minute background fluorescence signal, where N is a positive integer.

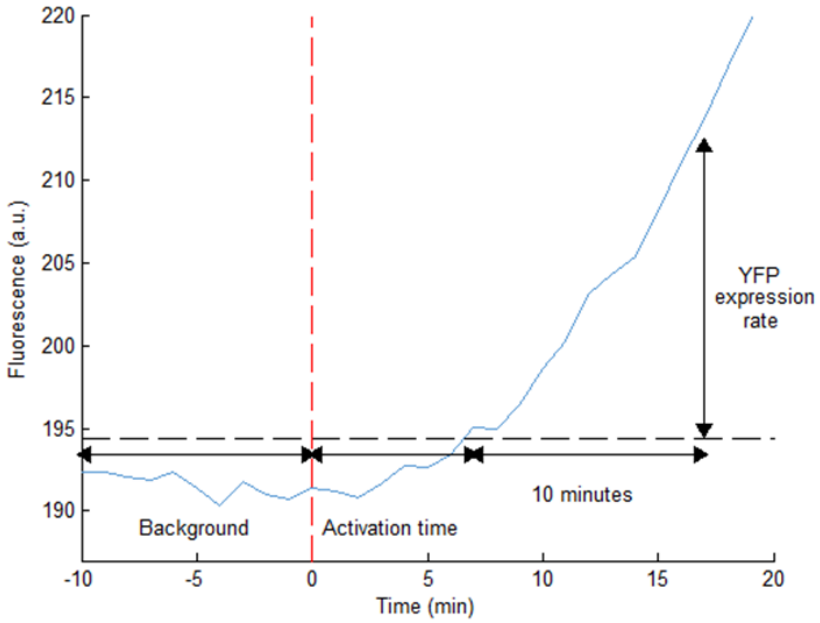

Figure 16. Estimation of activation time and YFP expression rate. The 10-minute background signals were used to define a threshold for estimating the activation time. The YFP expression rate was estimated by calculating the difference in YFP fluorescence 10 minutes after the activation time.

When N < 4, the estimated delay of some cells was shorter than the YFP maturation time (N=3 case, Figure 17), which is unrealistic. This is probably a result of using only 10 data points for reference. Therefore, the threshold is defined to be the mean plus four standard deviations of the 10-minute background fluorescence signal. The threshold was chosen to be minimal to reflect the timing of YFP increase, but not to underestimate the activation time.

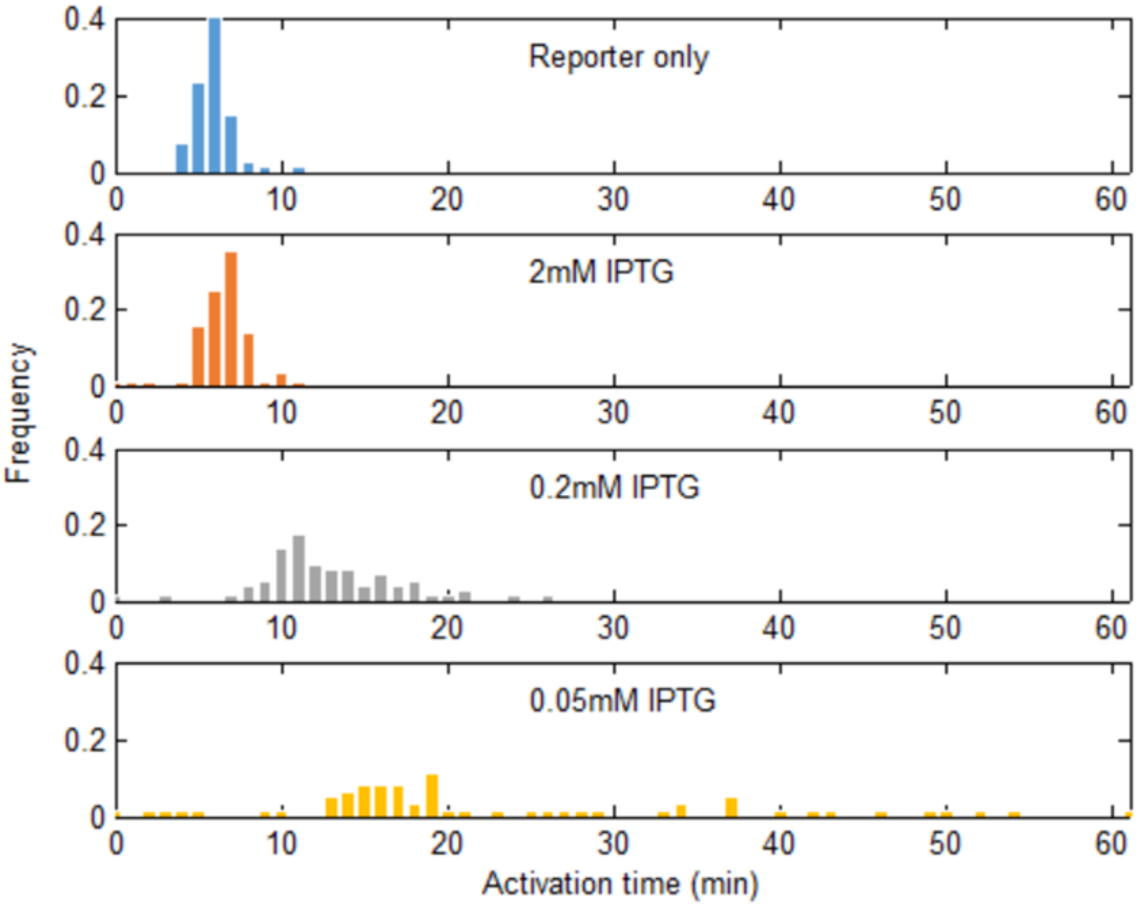

Figure 17. Underestimation of activation time when N = 3. Some cells are estimated to be activated within four minutes, which is shorter than the maturation time of the P_{BAD} reporter-only circuit.

The estimated YFP maturation time was 6.4 ± 1.3 minutes based on the measurement of the reporter itself (Figure 18). For the activation circuit, when the inducer concentration was high (2 mM IPTG), the estimated activation time was 7.2 ± 1.4 minutes. Subtracting the mean of YFP maturation time from the mean of activation time, the 2 mM IPTG data shows that the activation can take place in less than one minute.

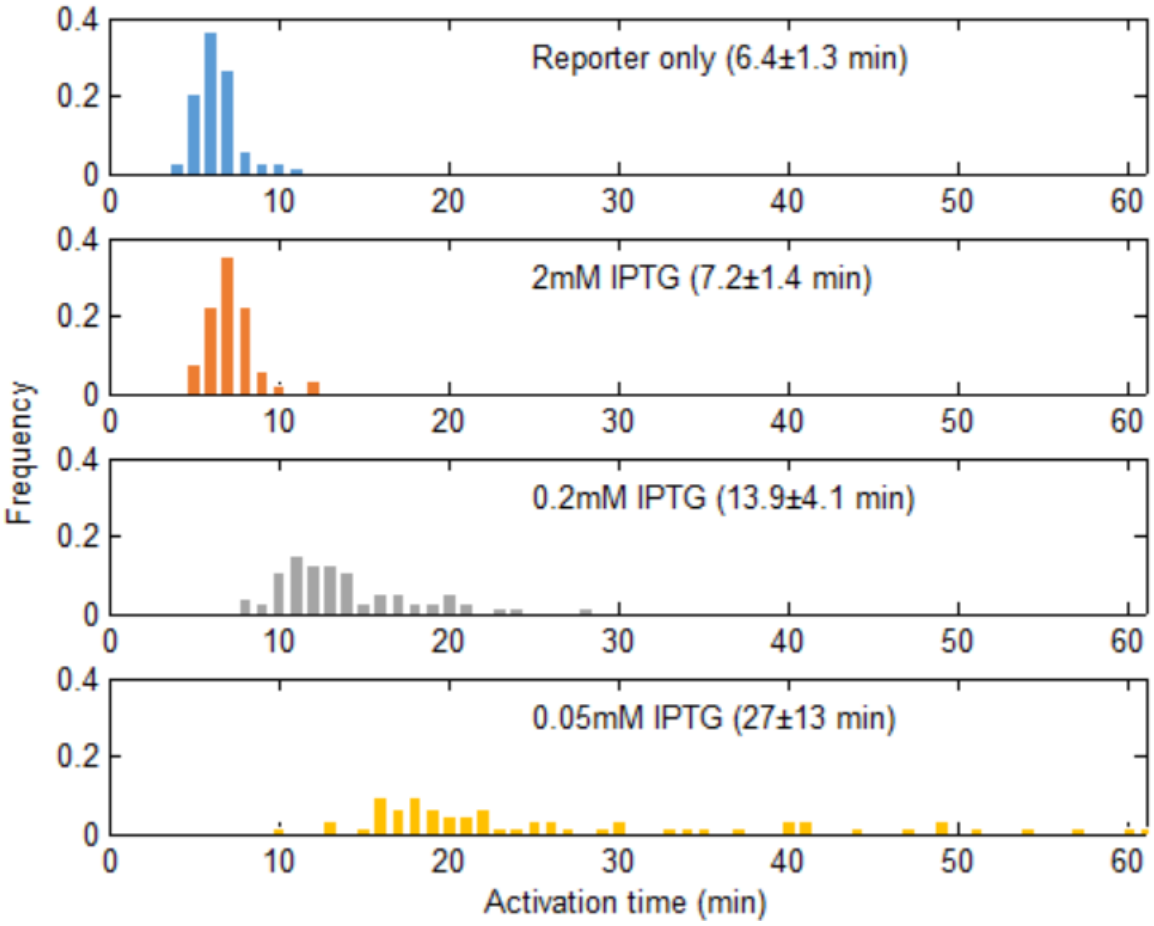

Figure 18. Estimated activation time from the activation circuit and the P_{BAD} reporter-only circuit. The activation time was only a few minutes and is close to the P_{BAD} reporter-only circuit when triggered with 2 mM IPTG. Reducing IPTG concentration resulted in longer activation time, and the time becoming variable among single cells.

Reducing the IPTG concentration leads to an increased delay; the activation time increased to 13.8 ± 4.1 minutes at 0.2 mM IPTG, and 27.9 ± 12.9 minute at 0.05 mM IPTG. The mean and variability both increased when IPTG was reduced. This shows that the production rate of AraC could affect the timing of P_{BAD} activation.

To quantify the YFP expression rate (ΔYFP), the fluorescence 10 minutes after the activation time was used (Figure 16). The expression of the activation circuit at

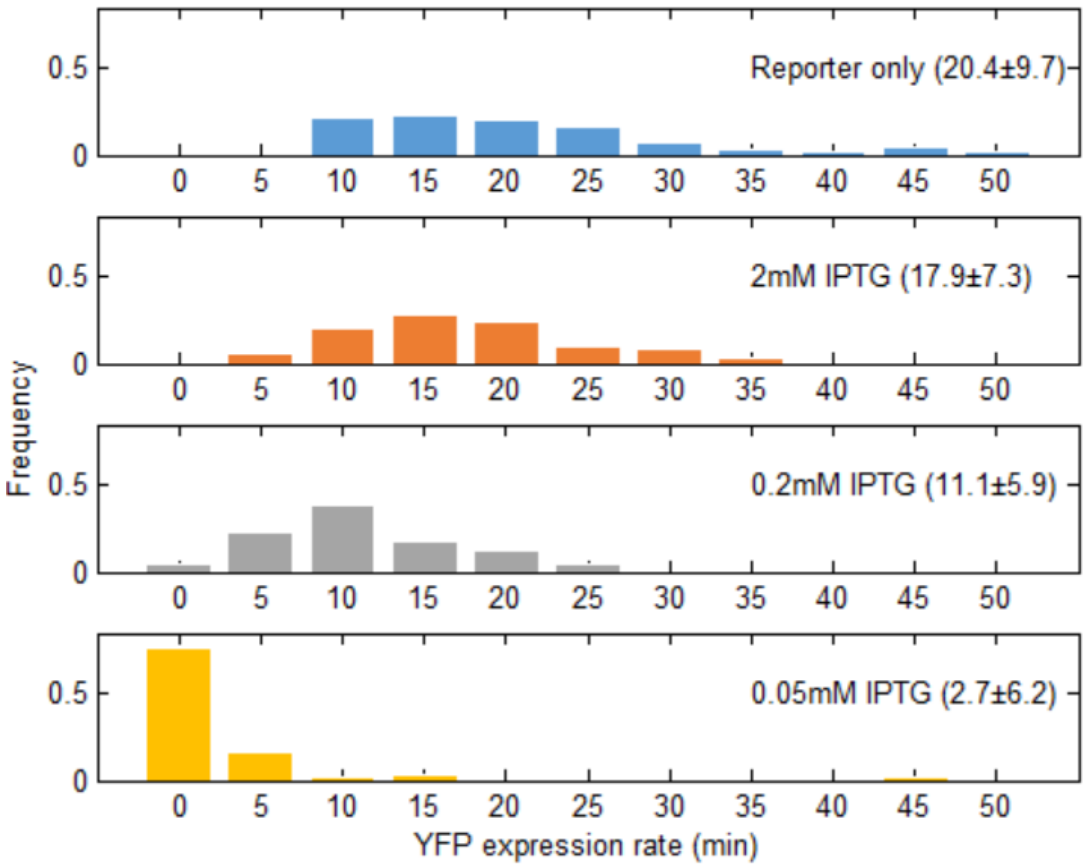

Figure 19. Estimated YFP expression rate from the activation circuit and the P_{BAD} reporter-only circuit. The YFP expression rate of activation circuit was close to the P_{BAD} reporter-only circuit when triggered with 2 mM IPTG. Reducing IPTG concentration resulted in a lower YFP expression rate.

2 mM IPTG was close to the P_{BAD}-only reporter circuit (Figure 19). The expression decreases when the inducer is reduced. At 0.05 mM IPTG, cells did not express much YFP, though the YFP signal did pass the defined threshold. This shows that AraC is not able to fully trigger P_{BAD} activation at low IPTG concentration.

Finally, a linear regression analysis for the YFP expression rate and the activation time was conducted (Figure 20). The results show that they are negatively correlated (r =

–0.3967, P = 0.0001 for 2 mM IPTG; r = –0.4857, P = 0.0001 for 0.2 mM IPTG; r = –0.2495, P = 0.0525 for 0.05 mM IPTG). These results indicate that the cells that are turned on faster have higher YFP expression. On the contrary, for cells that are turned on slowly, the YFP expression is low. Hence, the AraC expression might be correlated with YFP expression by extrinsic noise (Elowitz et al., 2002). The P_{BAD} reporter-only circuit showed no correlation between the activation time and YFP expression rate (correlation coefficient r = 0.0055, P = 0.9643).

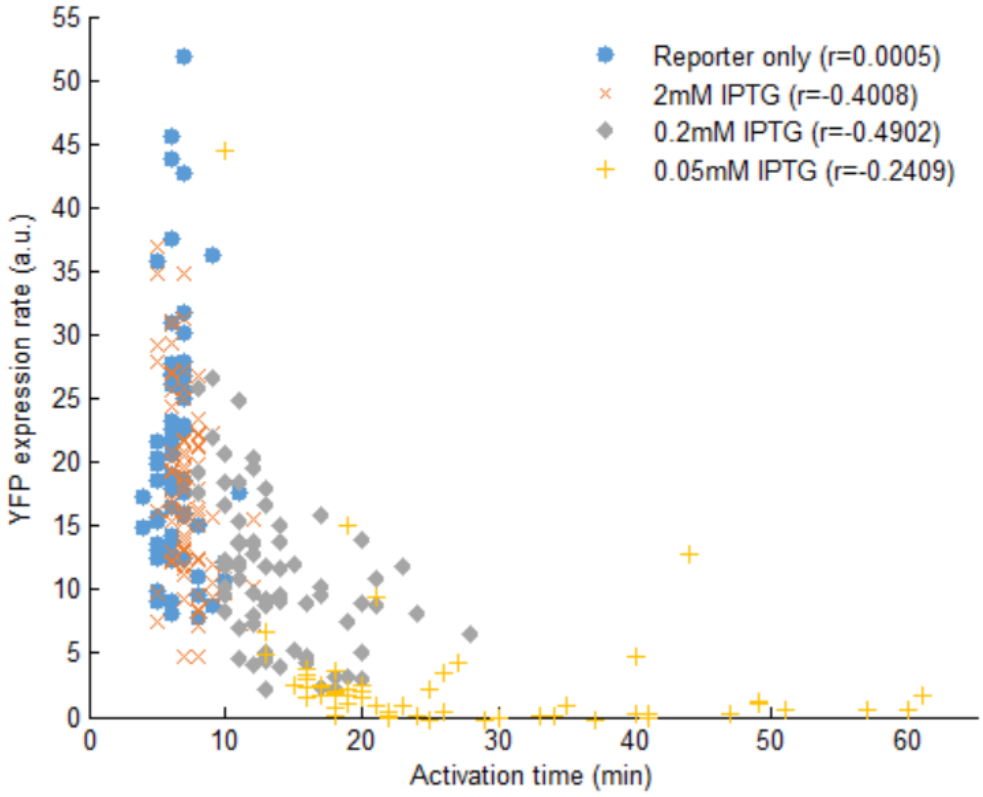

Figure 20. Scatter plot of YFP expression rate and activation time of single cells. The correlation coefficients showed that YFP expression rate and activation time are negatively correlated for the activation circuit for all IPTG concentrations tested. The two variables are not correlated for the P_{BAD} reporter-only circuit.

3.2. LacI repression time

The repression circuit was used to measure the transcriptional repression of $P_{lac/ara}$ promoter by LacI (Figure 21A). The LacI gene was placed under control of the inducible promoter $P_{BAD,}$ and the YFP gene was placed under control of the $P_{lac/ara}$ promoter. $P_{lac/ara}$ is a hybrid promoter, which is activated by AraC and repressed by LacI. Once ARA is added, LacI and YFP will be activated by the constitutively expressed AraC. Because of the transcriptional delay of LacI, the YFP signal will increase first, then decreases. By changing the ARA concentration, I was able to see how the production rate of LacI affected the transcriptional delay. The magnitude and duration of the pulse are determined by the promoter activities of P_{BAD} and $P_{lac/ara}$ under different ARA concentrations (Figure 21B).

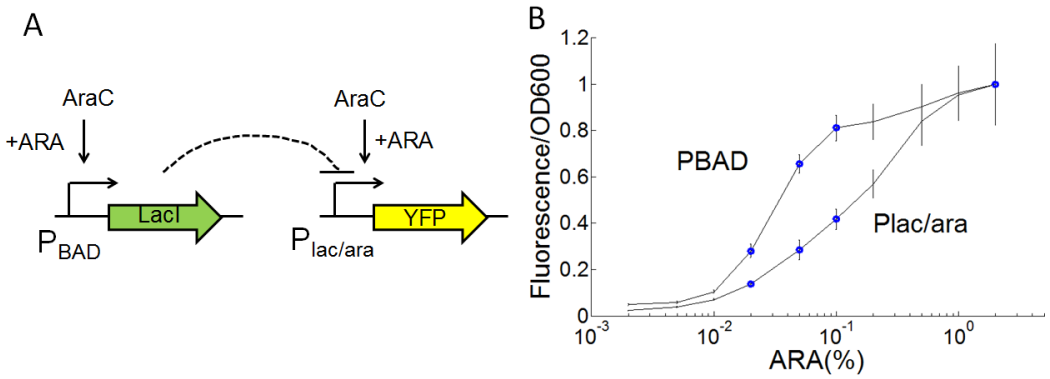

Figure 21. Measurement of repression delay. (A) The repression circuit. Adding ARA will trigger both the productions of LacI and YFP. Then LacI will repress the productuib of YFP. (B) Four ARA concentrations were tested on the circuit to see how the production rate of LacI would affect the activation delay.

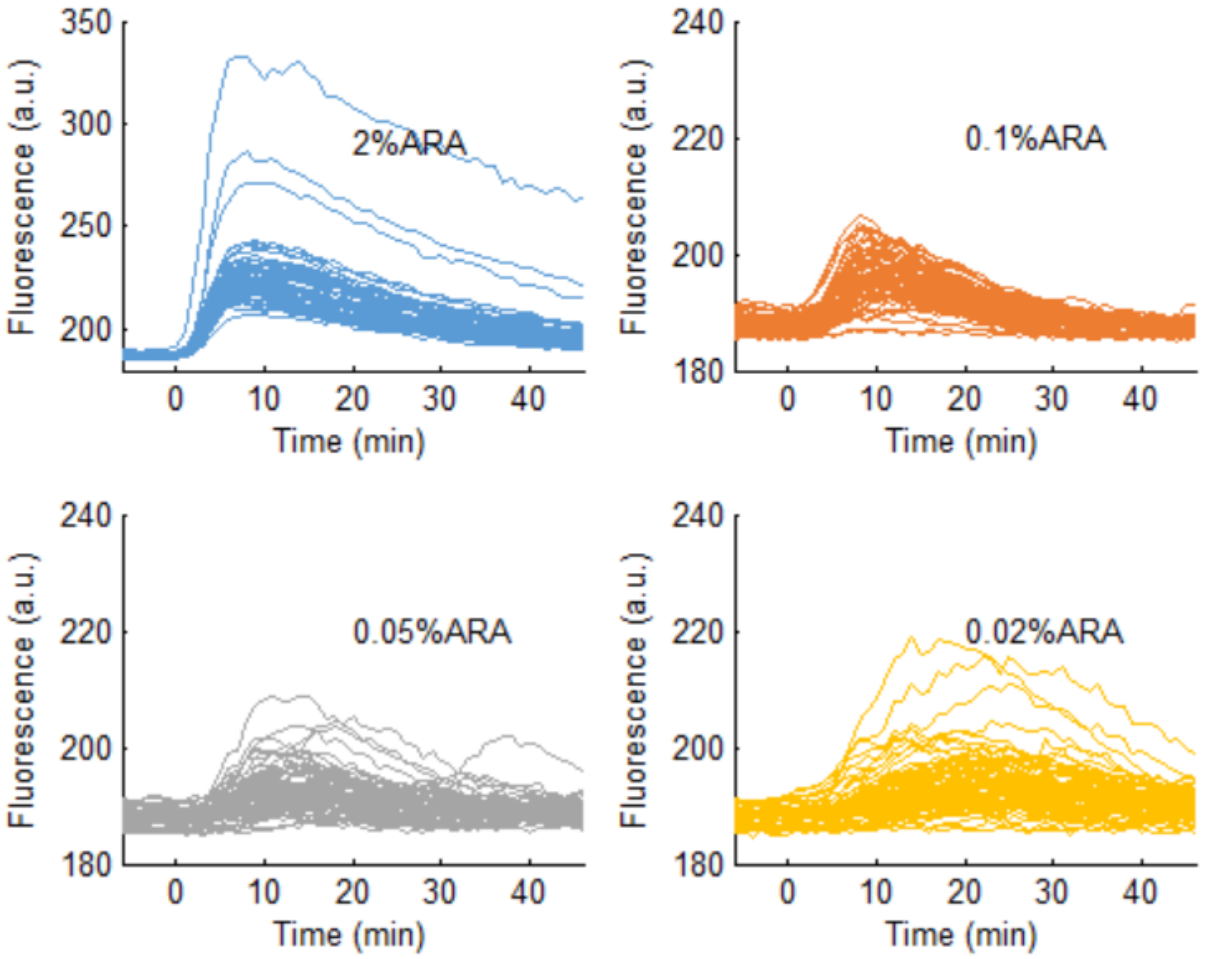

Figure 22. Single-cell fluorescence of repression circuit under 2%, 0.1%, 0.05%, and 0.02% ARA. Each experiment was repeated twice and at least 60 cells were collected.

The circuit was tested with four different ARA concentrations. The single-cell data shows pulses of YFP expression, with decreased amplitudes at lower ARA concentrations (Figure 22). In addition, the position of the peak of the pulse seems to appear later and be more variable.

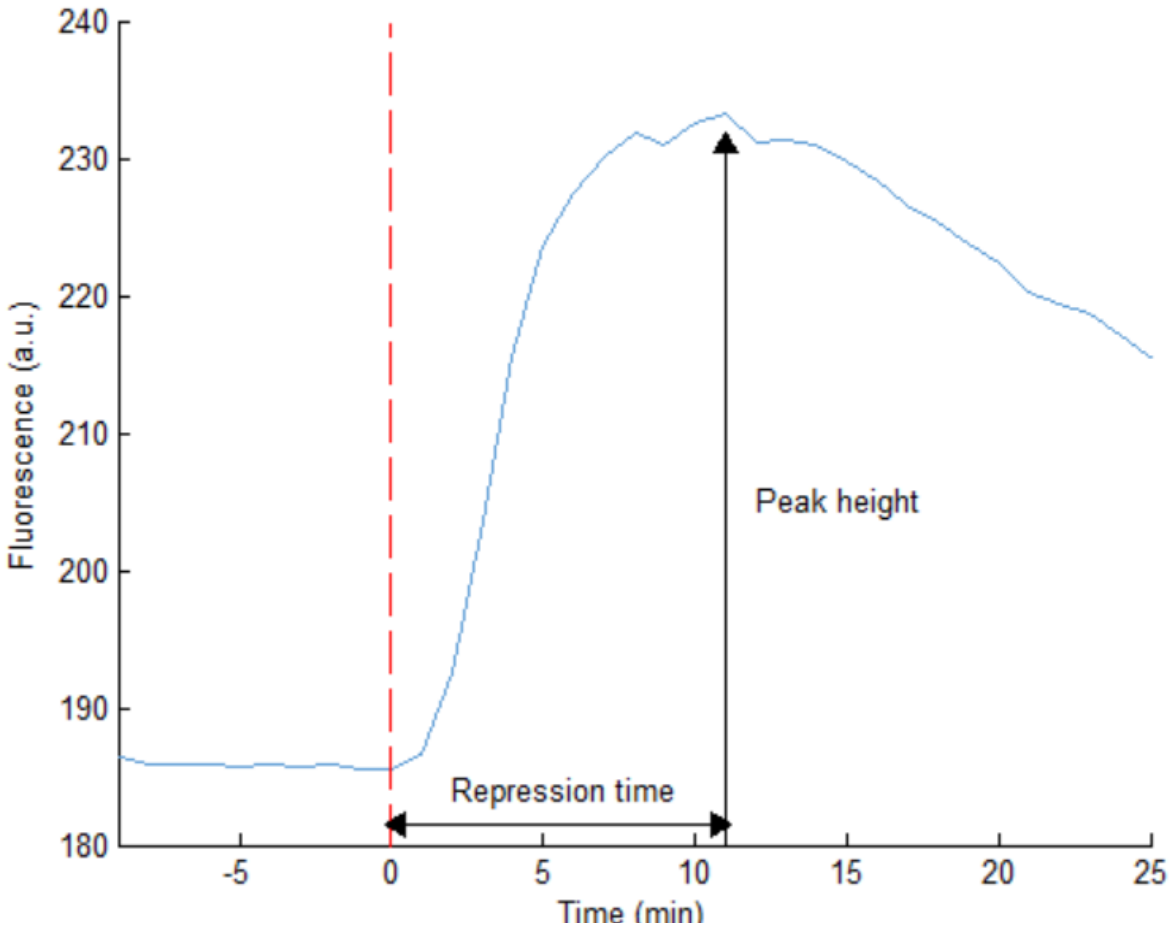

Figure 23. Estimation of repression time from the fluorescence pulse. The position of the peak was used to estimate the repression time. The peak height was also recorded.

To estimate the delay of transcriptional repression, the position of the peak is used to reflect the timing when $P_{lac/ara}$ is completely repressed by LacI (Figure 23). At 2% ARA, the repression time is 8.6 ± 1.1 minutes (Figure 24). The delay increases when ARA concentration is reduced: 10 ± 1.9 minutes at 0.1% ARA, 17.9 ± 9.4 minutes at 0.05% ARA, and 24.4 ± 9.7 minutes at 0.02% ARA. Combining the results of activation circuit, both data show that the delay increases when the TF production rate is reduced. The mean and variability increase.

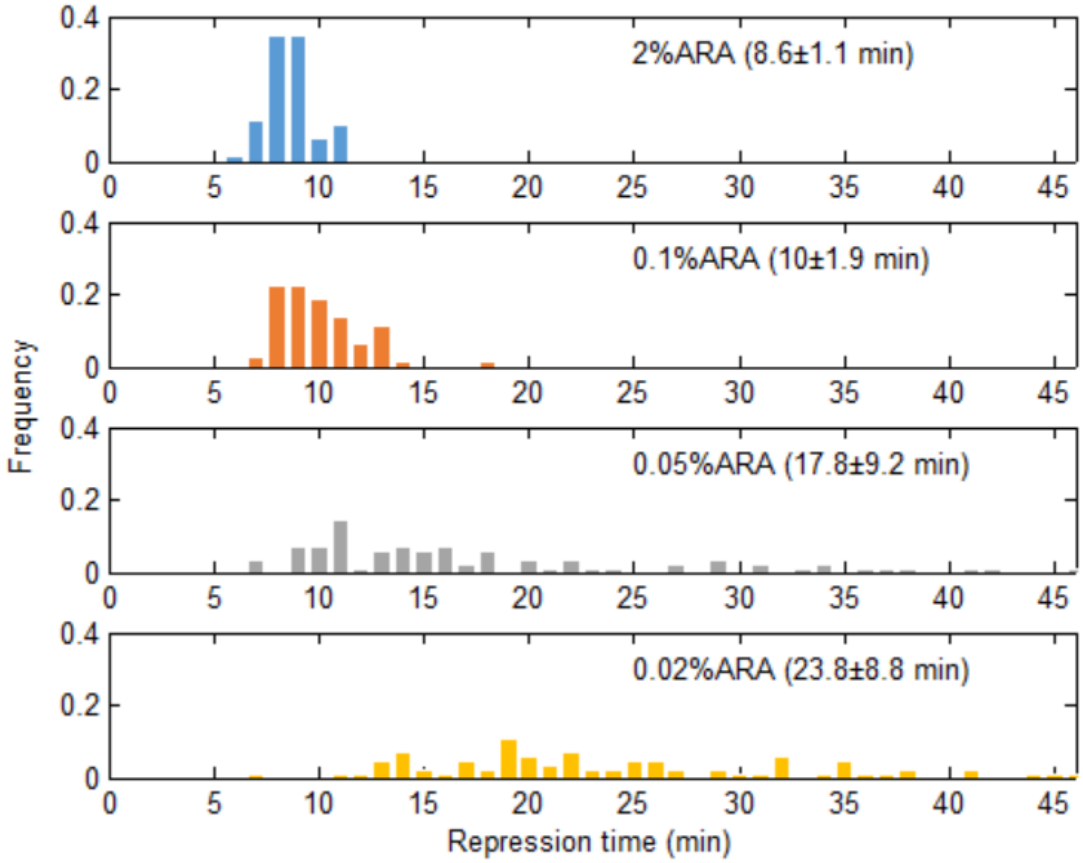

Figure 24. The estimated repression time from the repression circuit. The repression time was only a few minutes when triggered with 2% ARA. Reducing ARA concentration resulted in longer activation time and the time becoming variable among single cells.

Additionally, at very low inducer concentration, the mean of the delay is close to the cell doubling time (27.9 minutes for the activation circuit and 24.4 minutes for the repression circuit). Thus, the transcriptional delay can range from a few minutes to one cell doubling time, depending on the production rate of TF.

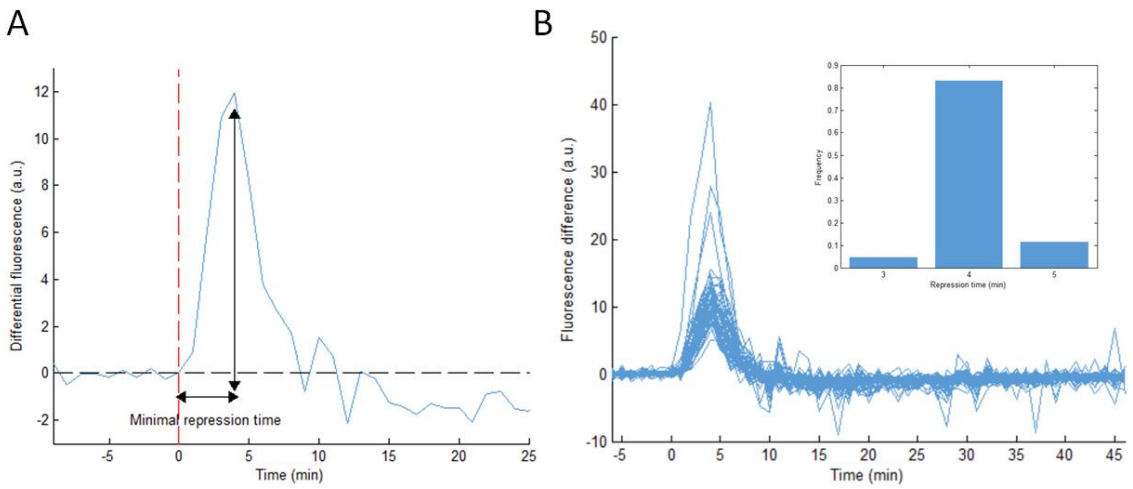

Figure 25. Estimation of minimal repression time of LacI. (A) For the fluorescence expression of the repression circuit at 2% ARA, the time difference was considered to obtain the differential fluorescence. The position of the peak reflects the timing of $P_{lac/ara}$ being affected by LacI and was used to estimate minimal repression time. (B) The estimated minimal time is 3–5 minutes.

Since AraC can activate P_{BAD} within one minute when it is triggered with 2 mM IPTG, it is of interest to know how quick LacI can become functional. The answer is important for modeling the robust genetic oscillator (Mather et al., 2009). To estimate it, I considered the time difference of the fluorescence signals at 2% ARA. The position of the peak indicates the time when the expression rate of YFP starts to decrease, so it reflects the timing when LacI becomes functional (Figure 25A). The results show that it takes 3–5 minutes for LacI to become functional, once it is triggered with 2% ARA (Figure 25B). The results are the same as the delay proposed by a theoretical model (Mather et al., 2009).

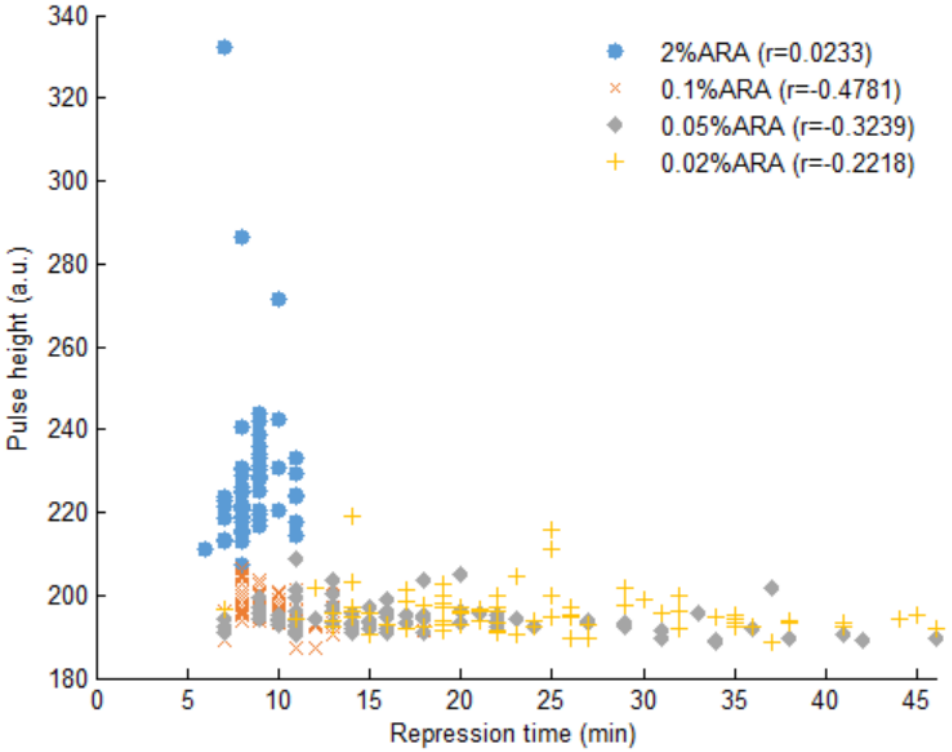

Figure 26. Scatter plot of peak height and repression time of single cells. The correlation coefficients showed that peak height and repression time are negatively correlated, except for 2% ARA.

Finally, I also did a linear regression analysis for the repression time and the height of the pulse (Figure 26). The results show that they are negatively correlated ($r = -0.4781$, $P = 0.0001$ for 0.1% ARA, $r = -0.3043$, $P = 0.0054$ for 0.05% ARA, and $r = -0.2642$, $P = 0.0165$ for 0.02% ARA). The exception is 2% ARA ($r = 0.0233$, $P = 0.8597$). This means that cells expressing more YFP tend to have shorter repression times. On the contrary, for cells that express less YFP, the repression time is longer. Hence, the LacI expression might be correlated with YFP expression. This is similar to the transcriptional activation.

3.3. Stochastic simulation for transcriptional regulation

To explain the observed phenomena for both the transcriptional activation and repression, I started with ODE models to fit the data. In the model, I only set variables for proteins, and every protein has two variables to represent the immature form and the functional form. The immature protein is used to account for the existing delay, such as the minimal transcriptional delay and YFP maturation time. The transition rate of the immature form to the functional form is based on the measured data. The production rate of TFs is scaled to the induction curve. The production rate of YFP is modeled as a Hill function (Alon, 2006). Proteins are diluted by cellular growth. If the protein is tagged with LAA, it is degraded by an enzymatic term. The model was fitted to the average of the single-cell fluorescence for all inducer concentrations tested. Then, the ODE model was transformed into the stochastic model following the Gillespie algorithm (Gillespie, 1977).

The activation circuit model has four variables: A, A*, Y, Y*, representing the concentrations of immature AraC, functional AraC, immature YFP, and mature YFP, respectively. The ODEs are as follows:

$$\frac{dA}{dt} = k_1(IPTG) - k_2 * A - \gamma_1 * A - \gamma_2 * \frac{A}{R + A + A^*}$$

$$\frac{dA^*}{dt} = k_2 * A - \gamma_1 * A^* - \gamma_2 * \frac{A^*}{R + A + A^*}$$

$$\frac{dY}{dt} = k_3 * \frac{(\frac{A^*}{\alpha})^\beta}{1 + (\frac{A^*}{\alpha})^\beta} - k_4 * Y - \gamma_1 * Y$$

$$\frac{dY^*}{dt} = k_4 * Y - \gamma_1 * Y^*$$

I fitted the parameters manually and with the fitted parameters:

$$k_1([2\text{mM IPTG}]) = 0.93 \text{ molecules}/(\text{min} * \text{cell volume})$$

$$k_1([0.2\text{mM IPTG}]) = 0.71 \text{ molecules}/(\text{min} * \text{cell volume})$$

$$k_1([0.05\text{mM IPTG}]) = 0.17 \text{ molecules}/(\text{min} * \text{cell volume})$$

$$k_2 = 1 \text{ min}^{-1}$$

k_3=10 molecules/(min $*$ cell volume)

$$k_4 = 0.5 \text{ min}^{-1}$$

$$\gamma_1 = 0.027 \text{ min}^{-1}$$

γ_2=1 molecules/(min $*$ cell volume)

$$\alpha = 15 \text{ molecules/cell volume}$$

$$\beta = 2$$

$$R = 30 \text{ molecules/cell volume}$$

The numerical simulation with the ODE model was close to the average fluorescence (Figure 27).

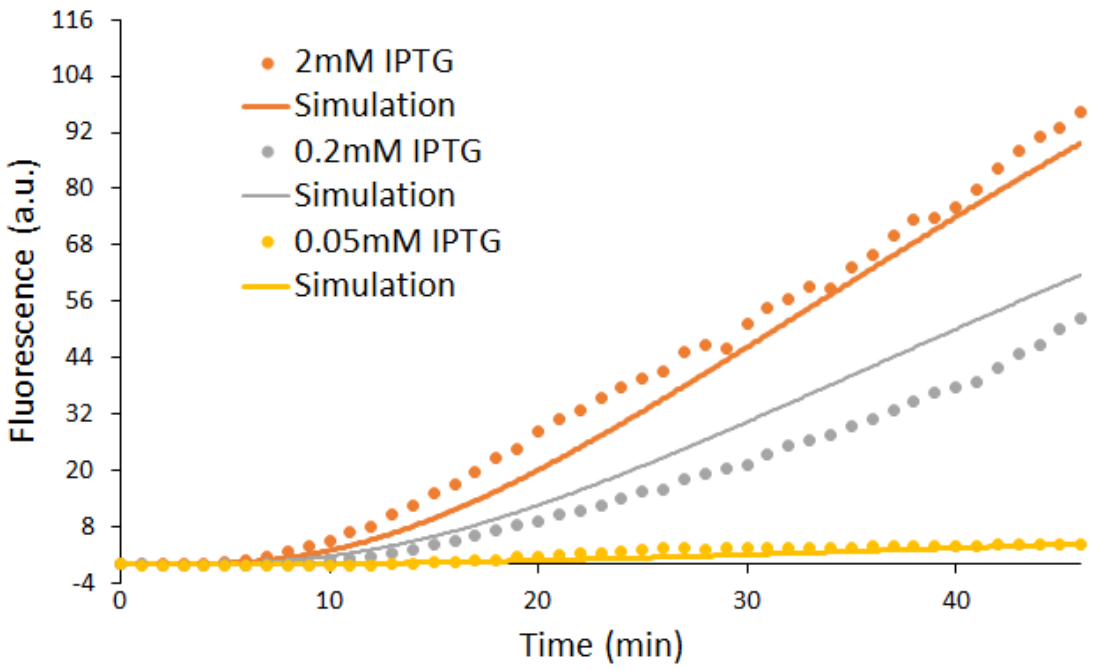

Figure 27. Numerical simulation of the ODE model of the activation circuit. The parameters of the ODE model were fitted to the mean fluorescence of single cells.

To transform the ODE model into the stochastic model, the rate terms were written as simple chemical reactions:

$$\emptyset \xrightarrow{k_1(IPTG)} A$$

$$A \xrightarrow{\gamma_1} \emptyset$$

$$A \xrightarrow{\gamma_2 * \frac{A}{R+A+A^*}} \emptyset$$

$$A \xrightarrow{k_2} A^*$$

$$A^* \xrightarrow{\gamma_1} \emptyset$$

$$A^* \xrightarrow{\gamma_2 * \frac{A^*}{R+A+A^*}} \emptyset$$

$$\emptyset \xrightarrow{k_3 * \frac{(\frac{A^*}{\alpha})^\beta}{1+(\frac{A^*}{\alpha})^\beta}} Y$$

$$Y \xrightarrow{\gamma_1} \emptyset$$

$$Y \xrightarrow{k_4} Y^*$$

$$Y^* \xrightarrow{\gamma_1} \emptyset$$

The model was simulated 1,000 times for the analysis using the C programming language. When the threshold for Y* was set at 1, the resulting activation time and YFP expression rate showed qualitatively similar results to experimental data (Figure 28).

For the repression circuit, the model has four variables: L, L*, Y, Y*, representing the concentrations of immature LacI, functional LacI, immature YFP, and mature YFP, respectively. The ODEs are as follows:

$$\frac{dL}{dt} = k_1(ARA)\text{-}k_2 * L - \gamma_1 * L - \gamma_2 * \frac{L}{R + L + L^* + Y + Y^*}$$

$$\frac{dL^*}{dt} = k_2 * L - \gamma_1 * L^* - \gamma_2 * \frac{L^*}{R + L + L^* + Y + Y^*}$$

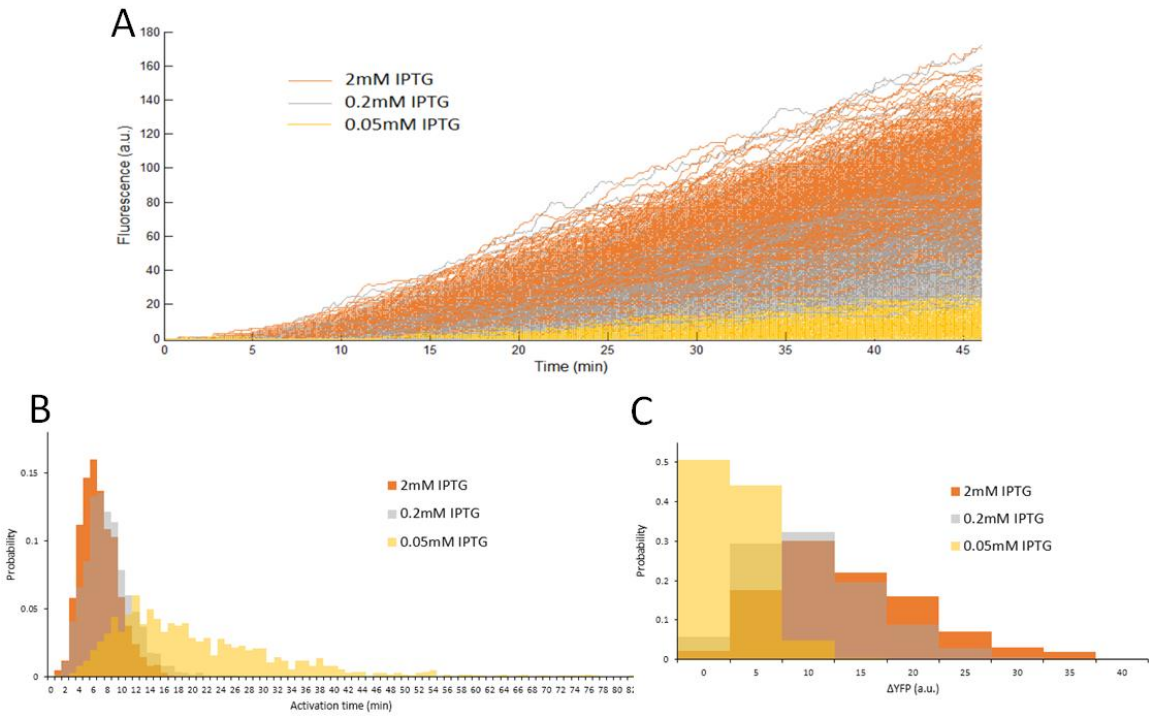

Figure 28. Numerical simulation of the stochastic model of the activation circuit. (A) 1,000 single-cell trajectories generated by the stochastic model. (B) Estimated activation time. (C) Estimated YFP expression rate.

$$\frac{dY}{dt} = k_3(ARA) * \frac{1}{1 + \left(\frac{L^*}{\delta}\right)^{\beta}} - k_4 * Y - \gamma_1 * Y - \gamma_2 * \frac{Y}{R + L + L^* + Y + Y^*}$$

$$\frac{dY^*}{dt} = k_4 * Y - \gamma_1 * Y^* - \gamma_2 * \frac{Y^*}{R + L + L^* + Y + Y^*}$$

With the fitted parameters

$$k_1([2\%\ ARA]) = 0.9\ \text{molecules}/(\text{min} * \text{cell volume})$$

$$k_1([0.1\%\ ARA]) = 0.85\ \text{molecules}/(\text{min} * \text{cell volume})$$

$$k_1([0.05\%\ ARA]) = 0.66\ \text{molecules}/(\text{min} * \text{cell volume})$$

$$k_1([0.02\%\ \text{ARA}]) = 0.28\ \text{molecules}/(\text{min} * \text{cell volume})$$

$$k_2 = 0.25\ \text{min}^{-1}$$

$$k_3([2\%\ \text{ARA}]) = 1.1\ \text{molecules}/(\text{min} * \text{cell volume})$$

$$k_3([0.1\%\ \text{ARA}]) = 0.35\ \text{molecules}/(\text{min} * \text{cell volume})$$

$$k_3([0.05\%\ \text{ARA}]) = 0.21\ \text{molecules}/(\text{min} * \text{cell volume})$$

$$k_3([0.02\%\ \text{ARA}]) = 0.14\ \text{molecules}/(\text{min} * \text{cell volume})$$

$$k_4 = 0.5\ \text{min}^{-1}$$

$$\gamma_1 = 0.027\ \text{min}^{-1}$$

$$\gamma_2 = 1\ \text{molecules}/(\text{min} * \text{cell volume})$$

$$\delta = 1.5\ \text{molecules/cell volume}$$

$$\beta = 2$$

$$R = 30\ \text{molecules/cell volume}$$

The numerical simulation of the ODE model was close to the average fluorescence (Figure 29).

To transform the ODE model into the stochastic model, the rate terms were written as simple chemical reactions:

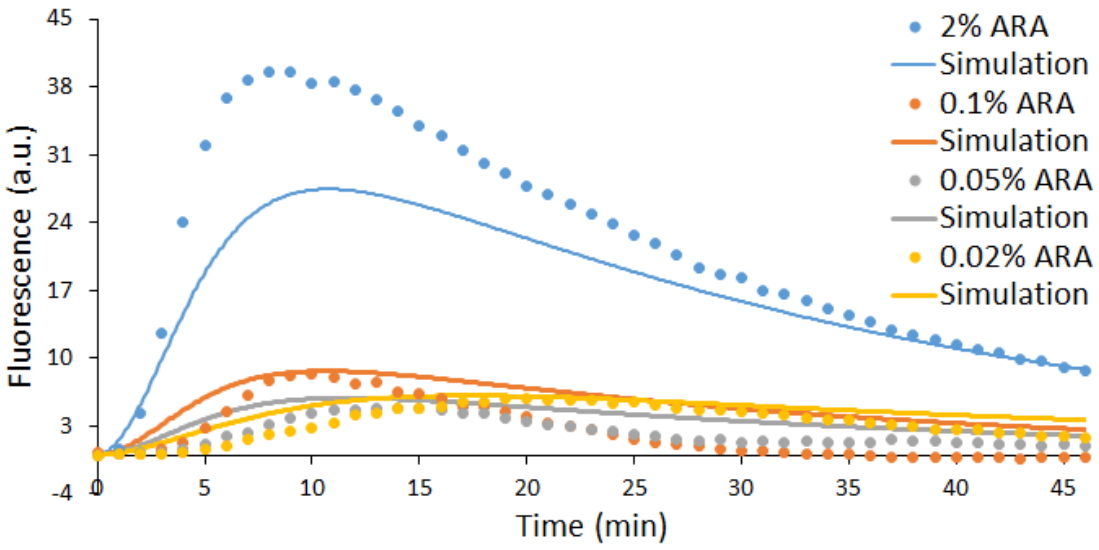

Figure 29. Numerical simulation of the ODE model of the repression circuit. The parameters of the ODE model were fitted to the mean fluorescence of single cells.

$$\emptyset \xrightarrow{k_1(ARA)} L$$

$$L \xrightarrow{\gamma_1} \emptyset$$

$$L \xrightarrow{\gamma_2 * \frac{L}{R+L+L^*+Y+Y^*}} \emptyset$$

$$L \xrightarrow{k_2} L^*$$

$$L^* \xrightarrow{\gamma_1} \emptyset$$

$$L^* \xrightarrow{\gamma_2 * \frac{L^*}{R+L+L^*+Y+Y^*}} \emptyset$$

$$\emptyset \xrightarrow{k_3(ARA) * \frac{1}{1+(\frac{L^*}{\delta})^{\beta}}} Y$$

$$Y \xrightarrow{\gamma_1} \emptyset$$

$$Y \xrightarrow{\gamma_2 * \frac{Y}{R+L+L^*+Y+Y^*}} \emptyset$$

$$Y \xrightarrow{k_4} Y^*$$

$$Y^* \xrightarrow{\gamma_1} \emptyset$$

$$Y^* \xrightarrow{\gamma_2 * \frac{Y^*}{R+L+L^*+Y+Y^*}} \emptyset$$

The model was simulated 1,000 times for analysis. The resulting activation time and YFP expression rate showed qualitatively similar results to the experimental data (Figure 30).

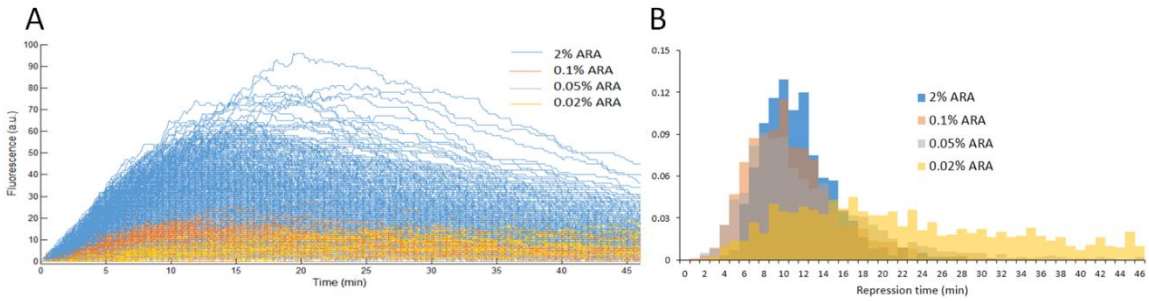

Figure 30. Numerical simulation of the stochastic model of the repression circuit. (A) 1,000 single-cell trajectories generated by the stochastic model. (B) Estimated repression time.

Here, I used stochastic models to explain the observed phenomena for both the transcriptional activation and repression. The increased delay and decreased YFP expression can be attributed to the existence of a threshold for TFs to regulate promoters. The increased variability of timing is also reflected in the stochastic model. Therefore, to model synthetic genetic circuits, the use of ODE model is sufficient. If necessary, small delays can be incorporated to reflect the minimal transcriptional delays. The negative correlation between delay and downstream gene expression, however, is not reflected in the model proposed. The results from the models showed that their correlation is positive (Figure 31). To account for this, we may need to incorporate sources of extrinsic noise, such as the partitioning of plasmids.

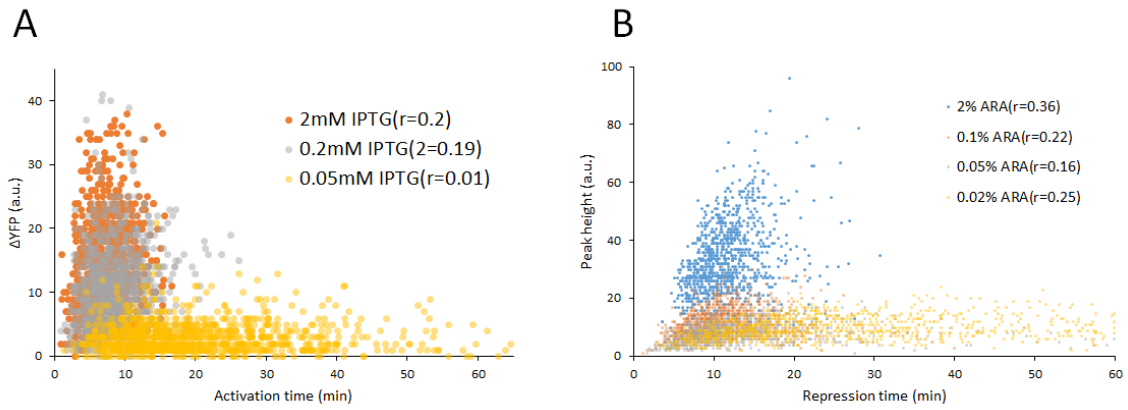

Figure 31. Correlation of delay and downstream gene expression from stochastic simulation. (A) Scatter plot of YFP expression rate and activation time. The correlation coefficients indicate that they are positively correlated. (B) Scatter plot of peak height and repression time. The correlation coefficients indicate that they are positively correlated.

3.4. Two-step genetic cascade and the convolved signaling times

The two-step genetic cascade was used to validate the measured delays of transcriptional regulation. Comparing the delay of the cascade with the convolved delays of single steps would give an estimate of the precision of the measurements. To build a two-step genetic cascade, the P_{lac} promoter of the AraC gene was replaced with the P_{Tet} promoter. The P_{Tet}-AraC circuit was combined with the repression circuit to complete the two-step genetic cascade (Figure 32A). When triggered with anhydrotetracycline (aTc) and ARA, AraC will be expressed first, and then it will activate both the LacI and YFP genes. The peak position of the pulse will then be the summation of the delays of activation and repression. When the cascade was triggered with 1 μg/mL aTc and 2% ARA, single-cell fluorescence showed a pulse with a delay relative to the repression circuit (Figure 32B).

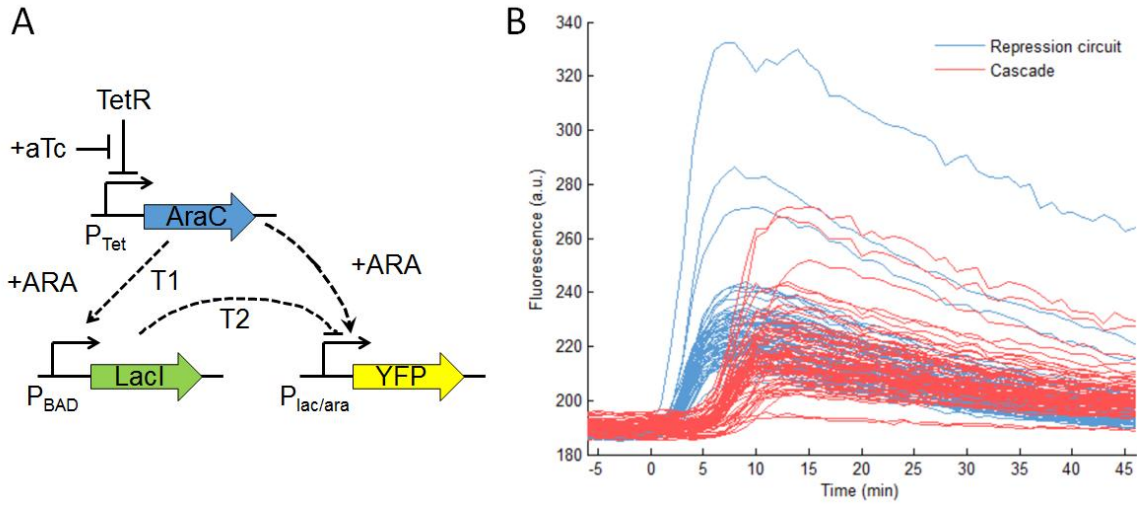

Figure 32. Single-cell fluorescence of the two-step genetic cascade and the repression circuit. After adding the activation step, the pulses of the cascade appear later relative to the repression circuit. Additionally, the heights of the pulses are smaller.

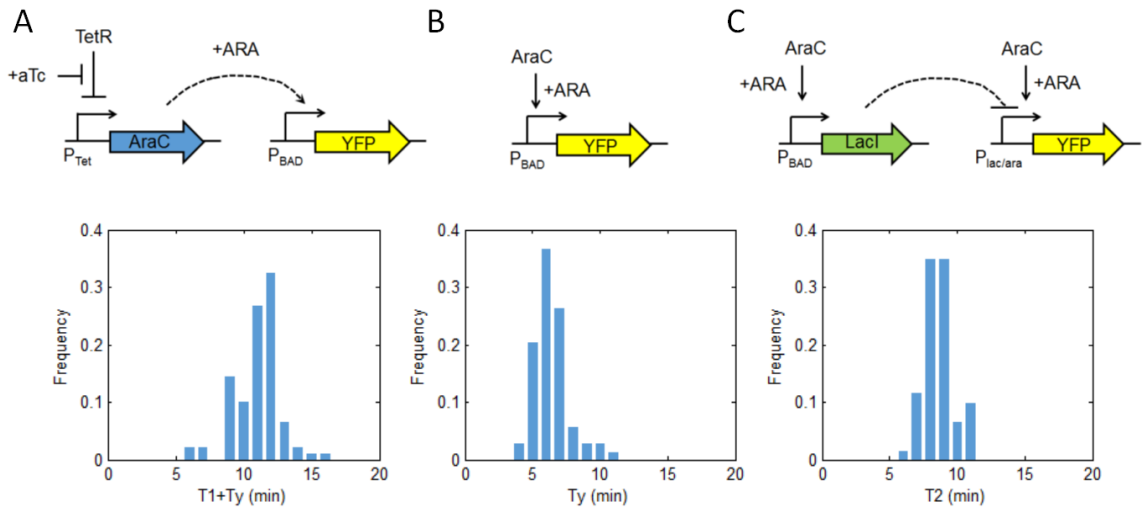

Figure 33. The constituents of the cascade and their measured delays. (A) The modified activation circuit was used to measure T_1+T_y. (B) The P_{BAD} reporter only circuit was used to measure T_y. (C) The repression circuit was used to measure T_2.

The position of the peak was used to determine the delay T_{1+2}. To measure the delays of single-step regulation, the modified activation circuit, repression circuit, and P_{BAD} reporter-only circuit were used (Figure 33). The modified activation circuit was triggered with 1 μg/mL aTc and 2% ARA. The measured activation time T_1+T_{YFP} was the summation of transcriptional activation delay T_1 and YFP maturation time T_{YFP}. The YFP maturation time T_{YFP} was measured with the P_{BAD} reporter-only circuit. The transcriptional repression delay T_2 was measured by the repression circuit.

To see whether the measured delay T_{1+2} of a two-step genetic cascade can be derived from the delays of its constituent steps T_1+T_{yfp}, $T_{yfp,}$ and T_2, the values for $T_1+T_2 = (T_1+T_{yfp}) - T_{yfp} + T_2$ needed to be calculated. These values were assumed to be independent random variables for performing the deconvolution and convolution. Numerically, this can be done with a Monte Carlo simulation:

1. T_i, T_j, and T_k are randomly selected from the data $T_1+T_{yfp} = T_{i1}, T_{i2}, \ldots, T_{iN1}$, $T_{yfp} = T_{j1}, T_{j2}, \ldots, T_{jN2}$, and $T_2 = T_{k1}, T_{k2}, \ldots, T_{kN3}$, where N1, N2, and N3 are the number of single cells measured.
2. Calculate $T_1+T_2 = T_i-T_j+T_k$. Store the value into an array.
3. Repeat step 1 and step 2 100,000 times.
4. Calculate the frequency of T_1+T_2.

The calculated T_1+T_2 is then compared to the measured T_{1+2}. The difference between the means of T_1+T_2 and T_{1+2} is of less than one minute (Figure 34A). However, the calculated standard deviation is slightly larger than the measured one. This might be a result of some correlations among the measured data. For example, if we assume some positive correlation between (T_1+T_{yfp}) and T_{yfp}, the mean T_1+T_2 is the same but the standard deviation would be smaller. The method used to calculate T_1+T_2 when (T_1+T_{yfp}) and T_{yfp} are positively correlated is as follows:

1. Assume (T_1+T_{yfp}) and T_{yfp} are normally distributed and fit the data with a normal distribution to calculate the means μ_{1+yfp} and μ_{yfp} and standard deviations σ_{1+yfp} and σ_{yfp}.

2. Use the property that the difference of (T_1+T_{yfp}) and T_{yfp} is still normally distributed. T_1 can be generated from the normal distribution with mean $\mu_{1+yfp} - \mu_{yfp}$ and standard deviation $\sqrt{{\sigma_{1+yfp}}^2 + {\sigma_{yfp}}^2 - 2\rho * \sigma_{1+yfp} * \sigma_{yfp}}$, where ρ is the correlation coefficient between (T_1+T_{yfp}) and T_{yfp}. Here I generated 100,000 points of T_1 from this distribution.

3. T_i and T_j are randomly selected from the data T_1=T_{i1}, T_{i2}, …, T_{iN1} and T_2=T_{j1}, T_{j2}, …, T_{jN2}, where N1 = 100,000, and N2 is the number of single cells measured.

4. Calculate T_1+T_2 =T_i+T_j. Store the value into an array.

5. Repeat step 3 and step 4 100,000 times.

6. Calculate the frequency of T_1+T_2.

Figure 34B shows the case for $\rho = 0.5$. The simulation shows that the mean is unchanged, but the standard deviation is closer to the measured value. Note that it is not the only way to get smaller standard deviation. If we assume that T_1 and T_2 are negatively correlated, we can get a similar result. To really understand what factors affect the convolution of delays in a multi-step gene cascade, we may fuse multiple fluorescent protein reporters to TFs, and characterize how the upstream regulator affects its target.

The single-step delays were convolved to calculate T_1+T_2 using the Monte Carlo method. When compared to T_{1+2} of the two-step genetic cascade, it shows that their means are very close, with a less than 1-minute difference. This demonstrates that the

measurements of single-step gene regulation are precise enough to reflect the summed delay of a two-step genetic cascade. The standard deviation of T_1+T_2 is larger than the standard deviation of T_{1+2}. This might be a result to small sample size. The single-delays might be also correlated. For example, if we assume that T_1+T_{YFP} and T_{yfp} are positively correlated, the standard deviation of T_1+T_2 can be smaller, with the mean being the same.

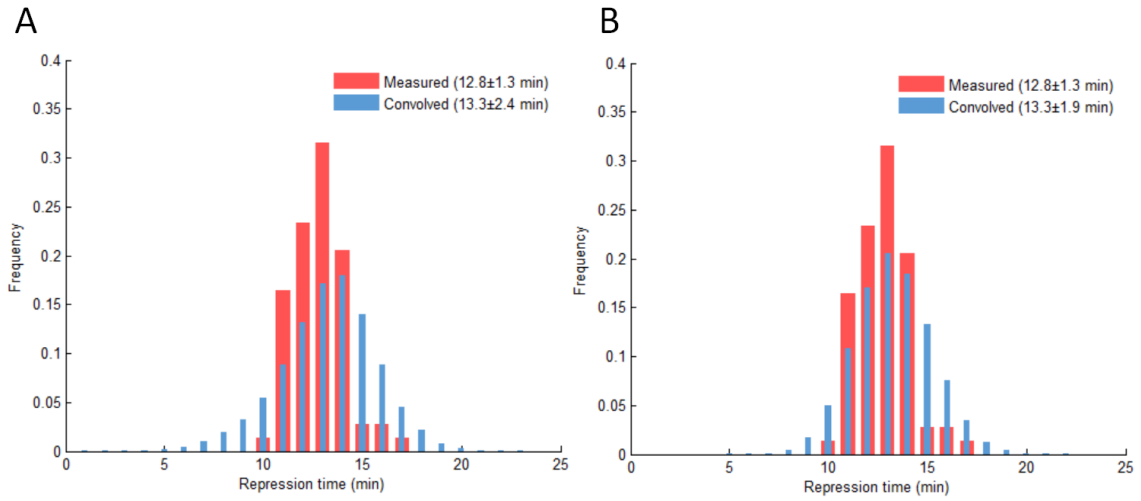

Figure 34. The measured delay of cascade and the convolved delay from single-step circuits. (A) The difference between the means of both delays is less than one minute. The variance of the convolved delay is larger than the measured one. (B) Reduced variance of convolved delay when assuming T_1+T_y and T_y are positively correlated with r = 0.5. The difference between the standard deviations is less than one minute.

3.5. Stoichiometry of signaling molecules

To gain more insight into the process of transcriptional regulation, I examined the known parameters of the signaling molecules, and compared our data to times predicted by those parameters. First, the YFP signal, Y, can be converted to numbers of single YFP molecules, N, per cell. I followed the same approach as described in the paper of

Rosenfeld and *et al* (Rosenfeld et al., 2005). The idea is that if YFP molecules are randomly segregated into daughter cells at cell division, the number of YFP molecules one daughter cell receives will follow a binomial distribution. To follow their approach, I first recorded the individual division events of the repression circuit at 2%, after the P_{lac} promoter was completely repressed (Figure 22). The events were grouped by the YFP signal strength. Then for each group, the root mean square (RMS) difference in fluorescence between two daughter cells was calculated, and fit to $\sqrt{\nu_y} * \frac{\sqrt{Y_1+Y_2}}{2}$, where ν_y is the fluorescence signal generated from one YFP molecule, and Y_1 and Y_2 are two daughter cells' fluorescence. The fitted ν_y is 0.14 fluorescence per YFP molecule. Then the number of YFP molecule can be calculated by $N = \nu_y * Y$ (Figure 35).

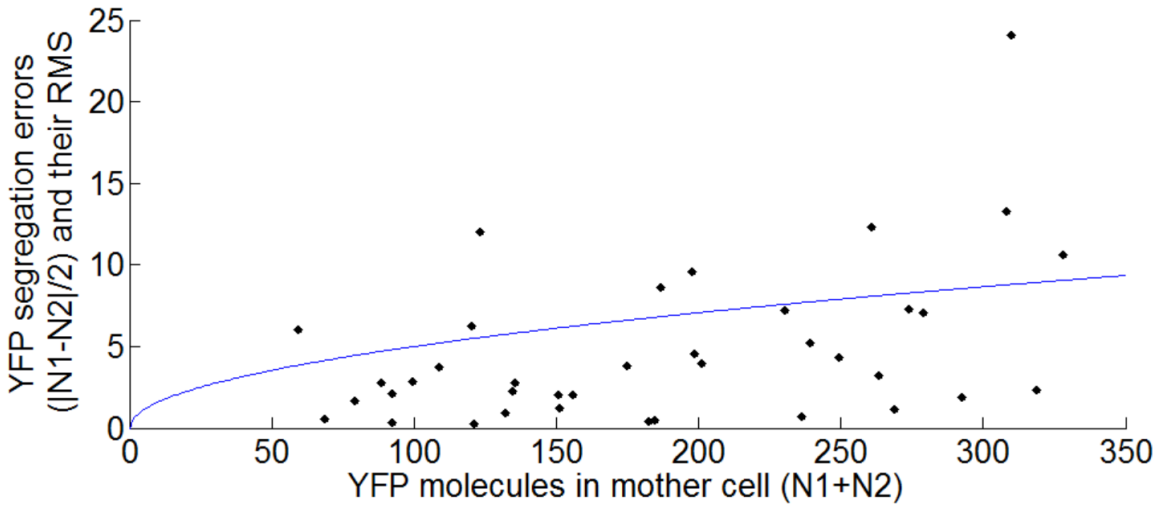

Figure 35. Binomial errors in YFP molecule partitioning at cell division. The flourescence signal difference of two daughter cells $\frac{|Y_1-Y_2|}{2}$ was used to estimate $\boldsymbol{\nu_y}$, which is the fluorescence generated from one YFP molecule. Then the flouresnce can be converted to numbers of YFP molecules $\boldsymbol{N}$ by $\boldsymbol{N = \nu_y * Y}$. Blue line is the fitted function $\frac{\sqrt{N_1+N_2}}{2}$.

Since the mean threshold for the determination of activation delay is 1.1 (Figure 15), it means that at least 1.1/0.14=8 YFP molecules are needed for the YFP signal to be higher than the background noise.

To make sure the measured delay corresponds to the known parameters of participated signaling molecules, here I examined the data of the repression circuit. The dissociation constant K_D of LacI is estimated to be 10 pM (Falcon and Matthews, 2000). Given that the size of an *E. coli* cell is roughly $1\ \mu m^3 = 1\ fL$, the unit of the dissociation constant can be converted to molecules per cell: $K_D = 10\ pM = 10\ \frac{pmole}{L} = 6 * 10^{12}\ \frac{molecules}{L} = 0.006\ \frac{molecules}{fL} = 0.006\ \frac{molecules}{cell}$. The dissociation constant corresponds to the LacI concentration at which half of the operator sites are repressed. Given the fact that there are three operator sites in the genome, the number is unrealistically small. Since the K_D was determined *in vitro*, the difference of the experiment environment, such as high salt concentration, may underestimate the dissociation constant of LacI. In order to calculate the *in vivo* dissociation constant, here I used a different approach. It was measured that on average one cell has 265 LacI monomers (Li et al., 2014), which means that at least 61 LacI tetramers are needed to repress the lac operator sites. From this we can assume that 31 LacI tetramers will give half repression, and the *in vivo* dissociation constant $K_D = 31 * 4\ \frac{molecules}{cell} = 123\ \frac{molecules}{cell}$.

Next, the LacI production rate can be estimated from the YFP production rate from the P_{BAD} promoter. The average YFP production rate from the P_{BAD} reporter only

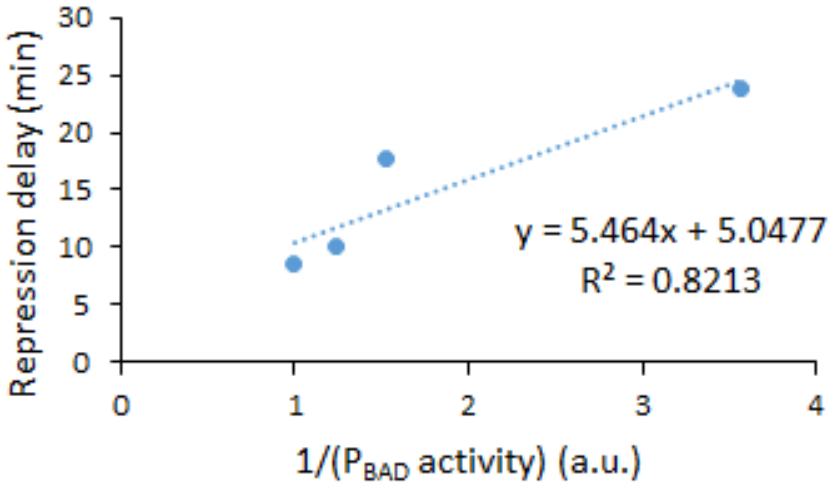

Figure 36. Repression delay is linearly proportional to the inverse of P_{BAD} promoter activity. The time delay for LacI to accumulate to a certain number is linearly proportional to its production rate, which directly depends on the P_{BAD} promoter activity.

circuit (Figure 15) is $\frac{2.5}{v_y}\frac{1}{min} = \frac{2.5}{0.14}\frac{molecules}{min} = 18\ \frac{molecules}{min}$. Assume that the protein production rates are same for YFP and LacI, when they are expressed from the P_{BAD} promoter, the production rate of LacI is $18\ \frac{molecules}{min}$. If we ignore the inducer uptake time, based on the measured delay of 8.6 minutes, it can be calculated that 18*8.6=155 LacI molecules are needed to completely repress the 20 lac operator sites on the plasmids. Then the dissociation constant K_D is roughly to be $\frac{155}{2}\frac{molecules}{cell} = 76\ \frac{molecules}{cell}$, which is close to the previous estimated $123\ \frac{molecules}{cell}$. Thus, the *in vivo* K_D of LacI estimated from these two approaches matches each other. At lower arabinose concentrations, the time needed to accumulate to 155 LacI is approximately proportional to the inverse of the P_{BAD} activity (Figure 36).

To examine whether the ligand uptake time can be ignored, here I estimate the transport rate of the ARA permease. I assume that the ligand uptake rate follows the Fick's law: $J = -D\frac{\partial C}{\partial z} \approx D\frac{\Delta C}{L}$, where D is the diffusion coefficient of the ARA, C is the concentration of the ARA outside cell, z is the direction perpendicular to the cell membrane, and L is the length of the permease. The diffusion coefficient D of the arabinose is $7.7 * 10^{-6}\ \frac{cm^2}{sec} = 7.7 * 10^{2}\ \frac{\mu m^2}{sec}$ (Nagy et al., 2009). For 2% arabinose, $\Delta C = 2 * 10^{-2}\ \frac{g}{mL} = 2 * 10^{-14}\ \frac{g}{\mu m^3} = 1.3 * 10^{-16}\ \frac{mole}{\mu m^3} = 7.8 * 10^{7}\ \frac{molecules}{\mu m^3}$. The regular size of a permease is $5\ \text{nm} = 5 * 10^{-3}\ \mu m$ (Li and Tooth, 1987). The transport rate is the flux multiplied by the cross-sectional area of the channel of the permease. The radius of the channel r is roughly $1\ \text{nm} = 1 * 10^{-3}\ \mu m$ (Pisponen et al., 2016). Based on these numbers, the transport rate of arabinose into cell is $J * \pi r^2 = D\frac{\Delta C}{L} * \pi r^2 = 7.7 * 10^{2}\ \frac{\mu m^2}{sec} * \frac{7.8*10^{7}\ \frac{molecules}{\mu m^3}}{5*10^{-3}\ \mu m} * 3.14 * 1 * 10^{-6}\ \mu m^2 = 3.8 * 10^{7}\ \frac{molecules}{sec}$. The dissociation constant of arabinose binding to AraC is $3 * 10^{-7}\ M = 1.8 * 10^{17}\ \frac{molecules}{L} = 1.8 * 10^{2}\ \frac{molecules}{fL} = 180\ \frac{molecules}{cell}$ (Phillips, 1976), which is the concentration at which half of AraC molecules are bound with the arabinose. Based on these numbers, the inducer uptake time is far less than in 1 second, even when the arabinose concentration is low. However, the estimation provided here is based on theoretical calculation and needs to be validated experimentally.

3.6. Conclusion

In summary, I measured the transcriptional delay of the transcriptional activator AraC and the repressor LacI by building synthetic genetic circuits. When TFs are highly induced, the delay can be as short as a few minutes. Reducing the expression of TFs leads to an increased delay. When the expression of TFs is very low, the mean of the delay is roughly the same as the doubling time of *E. coli*. In addition, the variability of the delay increases as the mean increases. This phenomenon can be explained by the existence of a concentration threshold for TFs to regulate the promoter (Figure 37A). The promoter activity can be modeled as a Hill function of TF concentration (Rosenfeld et al., 2005). To regulate the promoter, TF concentration needs to reach a threshold, H. When the TF is fully triggered, the time it takes for TFs to reach the threshold is short and the variability of the timing is small (Figure 37B). When the TF is weakly triggered, it takes longer to

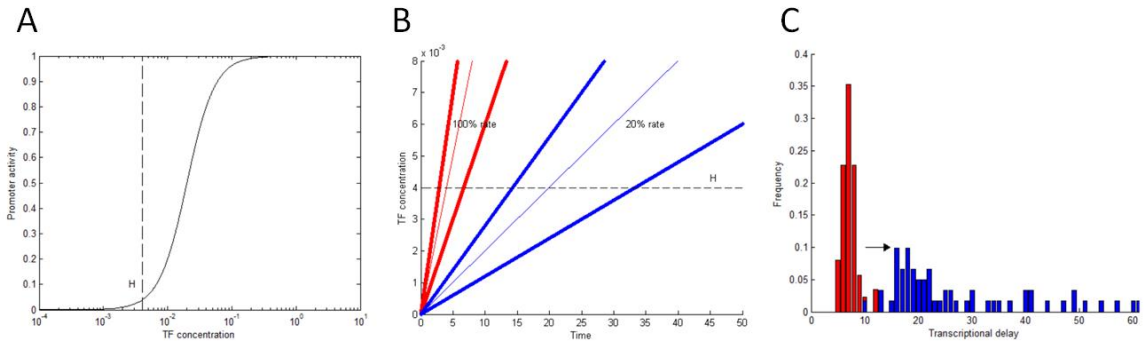

Figure 37. Increased delay can be explained by a threshold effect. (A) The promoter activity versus TF concentration. TF concentration needs to be above a threshold (H) to affect promoter activity. (B) TF concentration versus time when TF is produced at a constant rate. The 100% rate leads to a shorter delay than the 20% rate. The bold line shows the 50% fluctuation of TF production rate. The 20% rate tends to have larger variation of delay. (C) The measured activation time of AraC under 2 mM IPTG and 0.05 mM IPTG.

reach the threshold level and the variability of the timing becomes greater. Since the time is inversely proportional to the rate, the distribution of the timing will be right-skewed (Figure 37C). Furthermore, when the TF is weakly triggered, it does not regulate the promoter effectively. This can be seen in the YFP expression rate of the activation circuit. I also found that the transcriptional delay and the downstream gene expression are negatively correlated. This cannot be explained by the proposed stochastic model, which only incorporates the intrinsic noise of reactions. The negative correlation might be a result of extrinsic noise, which affects the TF and the reporter at the same time. For example, the plasmid copy number might be a source of extrinsic noise.

I further used the two-step cascade to show that the delay of cascade is close to the sum of single-step delays. The difference between the means is less than 1-minute. This shows that the measurements of the delay are precise enough to predict the timing of a more complex circuit. The standard deviation, however, is smaller than the summed delays. The difference in the variations of the delay distributions is possibly a result of correlations between single-step delays, which are not measured in this study. For example, plasmid copy number variation might lead to correlation between delays. If cells have more plasmids, they could express more TFs and YFPs, which will lead to shorter activation delay and YFP maturation time. When activation delay and YFP maturation time are positively correlated, the variation of the summed delay will be smaller. Also, resource competition could lead to reduced variability of the summed delay. Tabor *et al* designed competing RNA with the same RBS as a GFP reporter to study how the translational capacity affects gene expression (Tabor et al., 2008). They found that when the competing RNA is expressed, the mean GFP expression is lower and

it has larger variation. Therefore, one gene's expression could affect other genes. If cells express more AraC molecules, they may have less LacI molecules. Therefore, the activation delay and repression delay could be negatively correlated, and this will also lead to reduced variation of the summed delay.

The stoichiometries of the molecules were also examined. First, I found that my experiment setup can detect as low as eight YFP molecules, after converting fluorescence signal to the YFP molecules. Second, I found that the *in vitro* dissociation constant of the LacI repressor is underestimated when compared to the *in vivo* experiments. The copy number of LacI per cell was used to estimate the *in vivo* dissociation constant of LacI, which agrees with the estimation from the measured transcriptional delay of LacI.

Finally, I found that the time for AraC and LacI to be functional can be as short as a few minutes. This suggests that modeling synthetic genetic circuits only needs to incorporate a small constant delay rather than a distributed one. The longer delay at low inducer concentration can be explained by the existence of a threshold for TF to pass, which do not need to assume a long delay in the model. The incorporation of a small constant delay is especially important for gene circuits with feedback regulation, such as the oscillator.

3.7. Discussion

One immediate question from the results is whether other TFs show similar results and, if so, why. If TFs behave similarly, the same design principles can be applied to build similar functional modules. Then, as more genetic parts are being discovered

(Stanton et al., 2014), scaling-up synthetic genetic circuit design will become possible. However, why should they behave in the same way? Why is the delay bound by the cell doubling time? In the future study, we can test whether other TFs have similar delays. Since each TF has different binding affinity to the target promoter, the delay may be different. We could also characterize the delay for different copy-number plasmids or for genes on the genome to test how gene copy number affects the time delay. Other factors such as protein degradation rate, temperature, culture media could affect time delay as well. After a systematic study of the circuit's behavior, we will have a clearer picture about how the genetic circuit is working in the cell.

It is thought that the uptake of the inducer may contribute to most of the delay of the transcriptional regulation. The delay of inducer uptake at low concentrations has been shown to be long (Megerle et al., 2008). Here I used the fluorescence of the repression circuit at 0.02% ARA to examine the inducer uptake time (Figure 38A). The inducer uptake time is estimated by the activation delay as well, which is 9.7 ± 6.3 minutes (Figure 38B). Subtracting the inducer uptake time from the LacI repression time, we can estimate the delay from transcriptional regulation, which is 14.1 ± 6.8 minutes (Figure 38C). It shows that when the production rate of TF is low, transcriptional delay can be significant, and the variability of the delay is also large. To avoid the inducer uptake time, the transporters for the inducer can be expressed from plasmids to increase its number. Khlebnikov et al. have shown that by doing so the gene expression from P_{BAD} promoter is more homogeneous over a wide range of arabinose (Khlebnikov et al., 2001). It is then important to study how much variability of delay is from the inducer uptake.

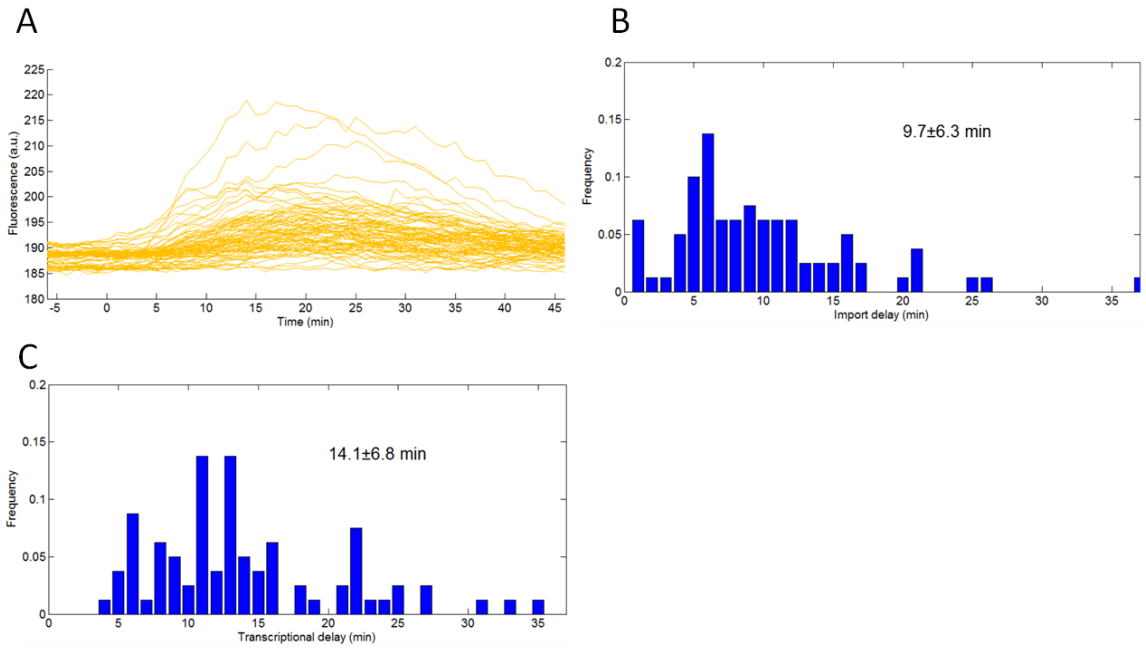

Figure 38. Import delay and transcriptional delay. (A) Single-cell fluorescence of the repression circuit at 0.02% ARA concentration. (B) The import delay is estimated by the activation time of the single-cell fluorescence. (C) The transcriptional delay is calculated by subtracting the activation time from the repression time in Figure 24.

To avoid the inducer uptake step and directly measure the transcriptional delay and, we can follow the optogenetics approach developed by Olson et al. (Olson et al., 2014). In their approach, the transcription factor is expressed from a light-inducible promoter, and the response time depends only on the phosphorylation step of the light-switchable proteins. Thus, we can avoid the possible long inducer uptake time at low inducer concentrations. Based on this optogenetics approach, they found that the delay for TetR to repress the P_{Tet} promoter is 7.0 ± 5.4 minutes, which is pretty similar to the measurement of LacI in this study (8.6 ± 1.1 minutes), although with a larger standard deviation. We could follow the same approach to repeat our measurements and see how the delay increases as the production rate of TF is decreased. With more data points, we

can see whether the delay is inversely proportional to the TF production rate. One drawback of the optogenetics approach is that it is not applied to the single-cell experiments of *E. coli* cells, which does not allow the characterization of the variability of delay. Also, the energy needed to excite the light-switchable proteins may not allow as frequent snapshots of cells as the microfluidic approach, which could lead to lower temporal resolution.

To understand how the delay is summed up in multi-step cascades, we can fuse a different fluorescent protein to the TF, and study how the TF concentration affects the output. This approach has been used to study how the noise from upstream gene expression is propagated to the downstream gene (Pedraza and van Oudenaarden, 2005), and to measure the gene regulation function (Rosenfeld et al., 2005). Then we can answer why the summed delay in the cascade has a smaller variability than expected. We can also use this approach to study how the delay affects circuit's dynamic behavior, especially for those with feedback regulations.

The transcriptional signaling time can be considered as a tunable variable in the synthetic genetic circuit design (Orosz et al., 2010). For example, it has been theoretically shown that increasing the transcriptional delay can stabilize bistable genetic circuits (Gupta et al., 2013). Since changing the inducer concentration could affect the delay, it is possible that modifying the genetic parts could affect the delay as well. For example, changing the RBS or promoter could affect the rate of the signaling process. Tools such as CRISPR interference (CRISPRi) (Qi et al., 2013) or antisense RNA (Hoynes-O'Connor and Moon, 2016) could also be used to change the delay. Then, we likely could fine-tune circuits' behavior, such as changing the frequency of the oscillator. The

tunability, however, is probably limited, since the effective delay can be no longer than one cell doubling time. Nevertheless, it is interesting to test the idea. For example, tuning the delay of negative feedback of the oscillator might lead to change in the period.

The timing of biological events is a largely unexplored area. It has been shown that increasing the length of a genetic cascade can improve the precision of timing (Amir et al., 2007). Theoretical works have also studied how the feedback could affect the precision of timing in gene expression (Ghusinga and Singh, 2016). With improved fluorescent protein reporters and microscopes, higher temporal resolution can be achieved. We can then study systems for which timing is important, such as cell differentiation (Narula et al., 2015) and cell cycle control (Murray, 1992). Also, how cells reduce the delay to deal with risks (Uphoff et al., 2016) and how they use delay as a bet-hedging strategy (Balaban et al., 2004) could be both interesting research topics.

Chapter 4

Antisense RNA-based gene regulation

Antisense RNA research is a side project of this focus, which aims to develop robust antisense RNA targeting genes. Unlike traditional antisense RNA approach, the antisense RNA design I propose here has the characteristic stem-loop structure to recognize the mRNA. The stem-loop structure is thermodynamically stable and will not be disrupted by RNA secondary structure. With 60bp antisense RNA, this approach successfully down-regulates the gene expression of four different proteins - YFP, mCherry, AraC, and LacI. The antisense RNA design proposed here can be applied to synthetic genetic circuit design.

4.1. Gene regulation methods for *E. coli*

Tools have been developed to control gene expression in bacteria. They allow researchers to easy turn on and turn off endogenous genes with encoding specific effector molecules. The most successful example is the CRISPRi method, which has been shown

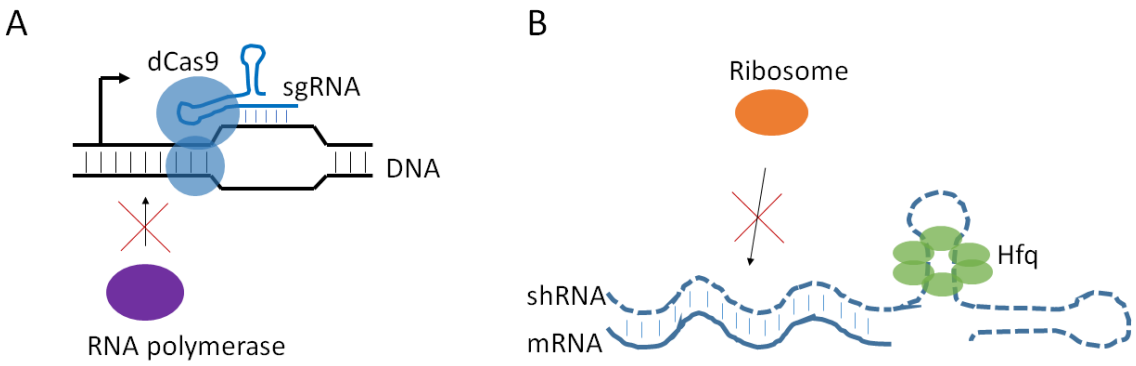

Figure 39. Two gene regulation methods used in *E. coli*. (A) The CRISPRi system is composed of sgRNA and dCas9. After binding to sgRNA, dCas9 will bind to the DNA sequences specified by the sgRNA. (B) shRNA is an antisense RNA with a short-hairpin structure. Protein Hfq will bind to shRNA and stabilize the binding betweeb mRNA and antisense RNA.

to down-regulate the genomically integrated YFP 1000 fold (Qi et al., 2013), and already been used to screen for essential genes in *Bacillus subtilis* (Peters et al., 2016). The CRISPRi system is comprised of two parts – the dCas9 protein and a short-guide RNA (sgRNA) (Figure 39A). The sgRNA contains a scaffold and a sequence complimentary to the target DNA. Once the dCas9 binds to the scaffold, the complex will bind to the DNA complimentary to the defined sequence. dCas will spatially block the proceeding of RNA polymerase and down-regulate the mRNA expression. Another commonly used gene regulation method is to use antisense RNA to bind to mRNA and disrupt the binding of ribosome. However, the repression efficiency of antisense RNA is unpredictable and oftentimes the antisense RNA design fails. For example, seven antisense RNAs were tested on RNaseE gene but only one antisense RNA succeeded (Kemmer and Neubauer, 2006). To improve the repression efficiency, it is suggested to add a short-hairpin RNA (shRNA) to the antisense RNA, which can stabilize the binding between mRNA and

antisense RNA (Na et al., 2013) (Figure 39B). The stabilization is due the binding of Hfq to the shRNA. Rather than introducing shRNA, the antisense RNA I proposed here uses a stable stem-loop structure to achieve higher efficiency. Next I will introduce the development of CRISPRi and antisense RNA methods.

4.1.1. CRISPR

CRISPR system is a prokaryotic immune mechanism that fights against the infection of phages or plasmids. Initially, it is found that prokaryotic genome contains repetitive sequences, which are flanked by spacers (Mojica et al., 2000). Since these repetitive sequences are homologous to the genes of invaders, it was postulated that these sequences may be used to recognize the invaders' DNA (Mojica et al., 2005). Then *cas* genes were discovered, which are located next to these repetitive sequences (Jansen et al., 2002). A study has found that the repetitive sequences and spacers are transcribed into CRISPR RNAs (crRNAs) in the presence of invaders (Marraffini, 2015). Then crRNAs were cleaved by the Cas proteins. Also, one Cas protein can bind to crRNA. Since crRNA contains sequences that are complimentary to the invading DNA, the complex can recognize and cut the invading DNA. It is estimated that approximately 40% of bacteria and 90% of archaea have CRISPR systems that target either foreign RNA (Hale et al., 2009) or DNA (Barrangou et al., 2007).

Among all kinds of CRISPR systems, the one from *Streptococcus pyogenes* is rather simple: it is only composed of one Cas protein (Cas9) and two RNAs (a crRNA and a trans-acting RNA (tracrRNA)) (Deltcheva et al., 2011). It was further found that crRNA and tracrRNA can be fused together to be one single-guide RNA (sgRNA), which

binds to Cas9 to cut the target DNA (Jinek et al., 2012). More importantly, by manipulating the sequences of sgRNA, this CRISPR system can be programed to target any DNA sequence for cleavage. After the DNA cleavage, cells will undergo either the non-homologous end joining or homology directed repair pathways. With a template DNA, the system can therefore be used to edit any genes (Hsu et al., 2014). The method has been applied to a broad range of organisms, including bacteria (Jiang et al., 2013), yeast (Dicarlo et al., 2013), worms (Waaijers et al., 2013), flies (Ren et al., 2013), fish (Hwang et al., 2013), plants (Xie and Yang, 2013), mice (Wang et al., 2013), and human cells (Cong et al., 2013).

Not only editing genes, CRISPR system can also be modified to control gene expression. After introducing point mutations to the nuclease domain, the Cas9 loses its nuclease function but still can bind to DNA. The deactivated Cas9 protein (dCas9) therefore can block the transcription elongation and down-regulate gene expression (Qi et al., 2013). The repression efficiency of this CRISPR interference (CRISPRi) method is high *E. coli*. However, when applied to mammalian cells, only modest repression was observed. The system was improved by fusing domains of transcription repressors to dCas9 (Gilbert et al., 2013). Other functional domains have been successfully fused to dCas9, such as transcription activator (Gilbert et al., 2013), chromatin remodeler (Keung et al., 2014) and fluorescent protein (Chen et al., 2013). Hence, CRISPR system can be as a modular platform for different applications.

4.1.2. Antisense RNA

Antisense RNA is a single-stranded RNA that is complimentary to mRNA. In the presence of antisense RNA, translation may be inhibited since mRNA could base-pair to it. Several antisense RNAs were discovered in E. coli, and they can be categorized into either *cis*-acting RNAs or *trans*-acting RNAs. *cis*-acting RNAs are usually located in the 5'- untranslated region (5' UTR) of mRNA. After transcription, *cis*-acting RNA will form a secondary structure and interfere with the binding of ribosome. In the presence of its complimentary RNA, the secondary structure will re-shape and the ribosome binding site can bind with ribosome again. The situation can be the opposite. The translation only proceeds in the presence of the complimentary RNA. An example of *cis*-acting RNA is the control of gene expression of RepC. RepC is a protein required for plasmid replication (Brantl and Wagner, 2002). Its expression is down-regulated by the antisense RNA, so the plasmid copy number is controlled. Unlike *cis*-acting RNA, *trans*-acting RNA is complimentary to mRNA, but not the 5' UTR. One example is the hok/sok system (Gerdes and Wagner, 2007; Gerdes et al., 1985). Hox is a toxin protein that kills cells. *sok* is the antisense RNA that binds to *hox* mRNA. The pair was found on R1 plasmid. After cell division, daughter cells which lose the plasmid will die.

Both *cis*-acting RNA and *trans*-acting RNA have been employed to control gene expression. For example, gene expression can be either activated or repressed by the antisense RNA by designing the 5' UTR (Chappell et al., 2015b; Lucks et al., 2011). The synthetic *cis*-acting RNAs have been shown to be highly specific and efficient (Green et al., 2014). Also, naturally occurring *trans*-acting RNAs have been compared to discover

consensus scaffolds (Na et al., 2013). It is thought that these scaffold will bind to some proteins and it could facilitate its binding with the mRNA. By fusing these scaffold to antisense RNA, the synthetic *trans*-acting RNA can be used to repress gene expression. However, there is no general design rules for both *cis*-acting RNA and *trans*-acting RNA, and often large library selection is required. Currently, CRISPR is more robust and efficient than the antisense RNA method.

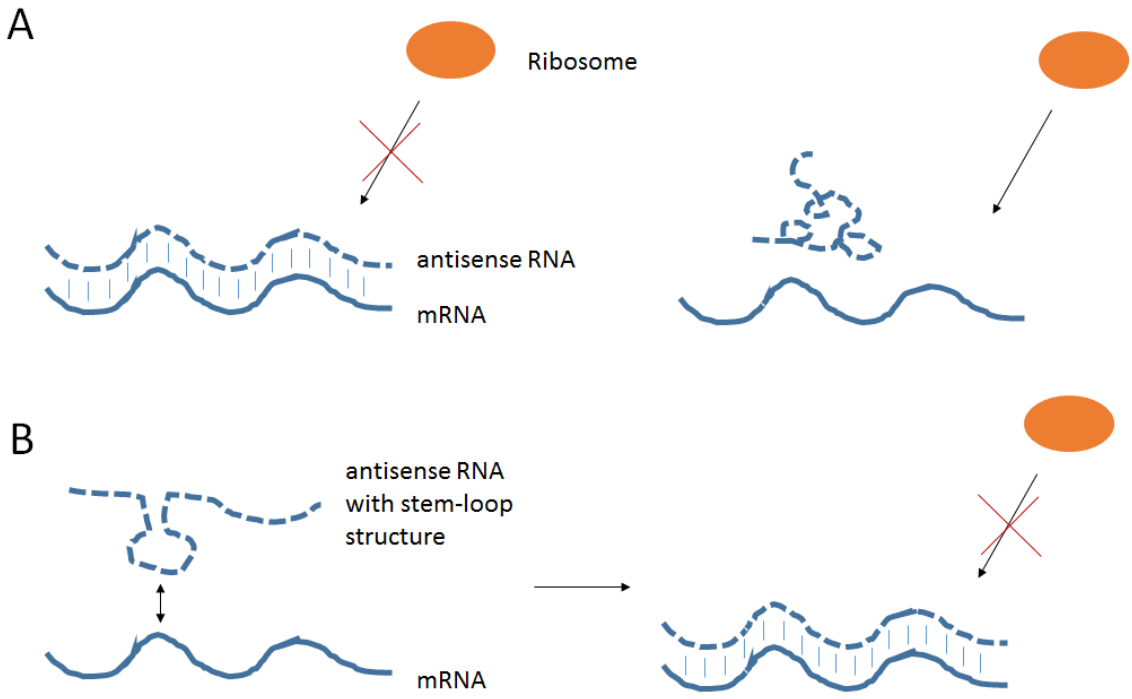

Figure 40. Antisense RNA design strategy. (A) Ideally, the antisense RNA can bind with mRNA and interfere with ribosome binding. The secondary structure makes antisense RNA bind to itself and to not be able to bind with mRNA. (B) The designed stem-loop structure is stable and can be used to recognize mRNA. Once the loop is bound, the stem will unwind and the remaining antisense RNA will bind with the mRNA.

4.2. Antisense RNA design

To solve the problem that the secondary structure of RNA will interfere the binding of antisense RNA to mRNA, I designed a stem-loop structure, which can prevent interference with the complimentary sequence (Figure 40). The strategy is to first find a non-folding sequence along the target sequences, and use it to construct the loop. Then the flanking sequences can be used to construct the stem. By choosing the appropriate length of flanking sequences, a stem-loop structure with sufficient stability can be designed. The loop then can be used to search the target RNA. When the loop of the antisense bind to mRNA, the binding energy will unwind the stem structure and the remaining antisense RNA can entirely bind to the mRNA. The procedure for designing antisense RNA is as follows:

7. Search for the longest sequence of the target gene that does not have secondary structure. The secondary structure can be calculated on the mFOLD website (http://unafold.rna.albany.edu/?q=mfold). Take YFP as an example, three 19bp sequences can be found that have no secondary structure (Figure 41). The most upstream one was used to construct the loop (tcaccttcaccctcgccac).

8. Starting from the non-folding sequence, select a 60bp sequence for the antisense RNA sequence (tcaccttcaccctcgccacgcacggaaaacttatgaccgttgacatcaccatccagttcc).

ATGCGTAAAGGCGAAGAGCTGTTCACTGGTGTCGTCCCTATTCTGGTGGAACTGG
ATGGTGATGTCAACGGTCATAAGTTTTCCGTGCGTGGCGAGGGTGAAGGTGACG
CAACTAATGGTAAACTGACGCTGAAGTTCATCTGTACTACTGGTAAACTGCCGGTA
CCTTGGCCGACTCTGGTAACGACGCTGACTTATGGTGTTCAGTGCTTTGCTCGTTA
TCCGGACCATATGAAGCAGCATGACTTCTTCAAGTCCGCCATGCCGGAAGGCTAT
GTGCAGGAACGCACGATTTCCTTTAAGGATGACGGCACGTACAAAACGCGTGCG
GAAGTGAAATTTGAAGGCGATACCCTGGTAAACCGCATTGAGCTGAAAGGCATT
GACTTTAAAGAAGACGGCAATATCCTGGGCCATAAGCTGGAATACAATTTTAACA
GCCACAATGTTTACATCACCGCCGATAAACAAAAAAATGGCATTAAAGCGAATTTT
AAAATTCGCCACAACGTGGAGGATGGCAGCGTGCAGCTGGCTGATCACTACCAG
CAAAACACTCCAATCGGTGATGGTCCTGTTCTGCTGCCAGACAATCACTATCTGAG
CTACCAAAGCGTTCTGTCTAAAGATCCGAACGAGAAACGCGATCATATGGTTCTG
CTGGAGTTCGTAACCGCAGCGGGCATCACGCATGGTATGGATGAACTGTACAAAA
CTAGTGCAGCGAACGACGAAAATTACGCCCTTGCAGCGTGA

Figure 41. The three non-folding sequences within YFP coding sequence. The non-folding sequence has no repeats and can be used to construct of the loop structure.

9. Add a 14bp sequence (taagttttccgtgc) to the 5’ end of the antisense RNA, which is complimentary to the downstream sequence of the non-folding sequence (gcacggaaaactta). The two 14bp flanking sequences will base-pair to form the stem structure (Figure 42).

10. PCR amplify the antisense RNA into a strong promoter such as $P_{LlacO-1}$ or J23100. Note that a transcription terminator is attached to the antisense RNA.

11. Co-transform the antisense RNA plasmid with the target gene plasmid. Induce the antisense RNA and measure the repression efficiency.

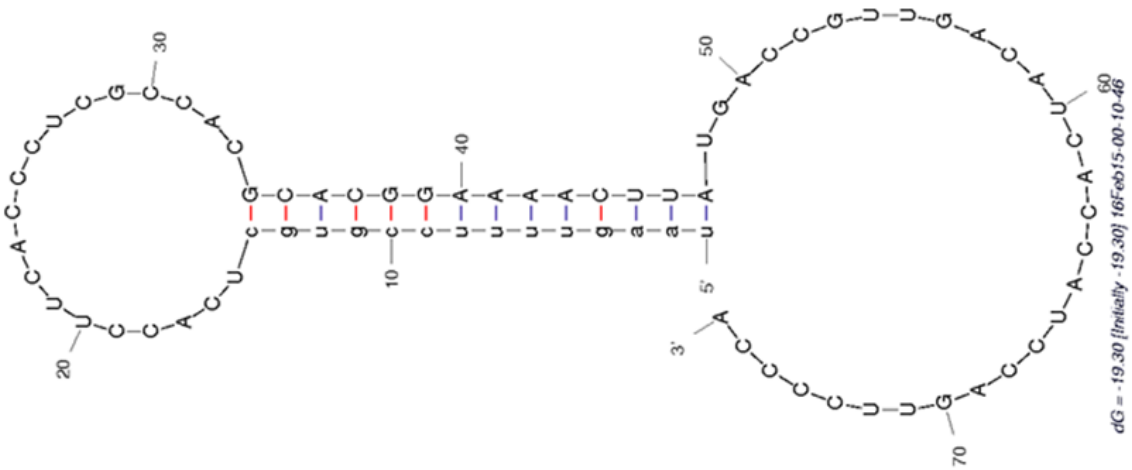

Figure 42. Secondary structure of antisense RNA for YFP repression. The stem-loop structure predicted from mFOLD is very stable.

4.3. Repression efficiency

To test the antisense RNA, the antisense RNA was expressed from strong promoters $P_{LlacO-1}$ and J23100 and the construct was on a high copy-number ColE1 plasmid. The target genes were on a medium copy-number p15A plasmid. The two plasmids were co-transformed into DH10BALT strain cells. After overnight culture, the sample was diluted 1:100 to LB and transferred to a 96-well plate, with the inducers added. After the plate was placed in the shaker at 37°C for two hours, the fluorescence and OD were read with an Infinite® 200 PRO fluorescence reader.

The antisense RNA was first tested on two fluorescent proteins - YFP and mCherry, which have quite different sequence compositions. The results showed that antisense RNA reduced 89% and 63% of expressions for YFP and mCherry, respectively (Figure 43). Next, the antisense RNA was tested on the transcription factors AraC and

LacI. The activation and repression circuits built previously were co-transformed with the antisense RNA plasmids. The results showed that antisense RNA reduced 44% and 29% of expressions for AraC and LacI, respectively.

Although the antisense RNA can reduce the expression of the chosen four proteins, it seems that the repression efficiency is not good on the TFs. A recent research showed a similar repression efficiency for the shRNA to repress TFs (Hoynes-O'Connor and Moon, 2016). Their results showed that the maximum repression is approximately 50%. The method I proposed here can be an alternative to other antisense RNA methods.

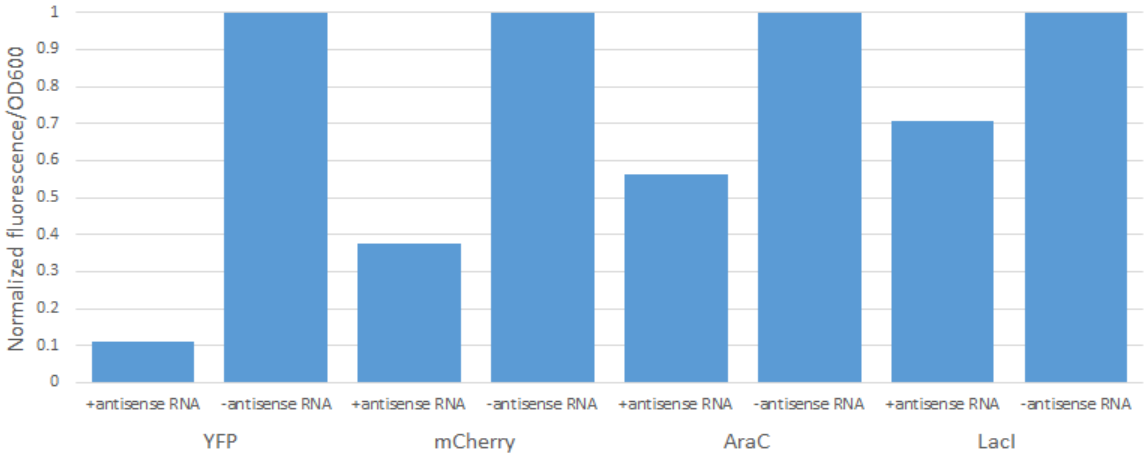

Figure 43. Antisense RNA efficiency on different genes. The antisense RNA works better for fluorescent proteins YFP and mCherry than transcription factors AraC and LacI.

4.4. Conclusion

Although the antisense RNA method is less efficient than the CRISPR method, it still has some advantages. First, the size of antisense RNA (~100bp) is much smaller than the CRISPR system (~4000bp). It can be easily implemented into existing designs. Also, the antisense RNA should be less toxic to CRISPR. The antisense RNA may be used to fine-tune the behaviors of existing synthetic genetic circuits. For example, change the period of the oscillator or reduce the leaky expression. It can also be used to repress the sgRNA, which creates an additional control (Lee et al., 2016).

References

Alon, U. (2006). An Introduction to Systems Biology: Design Principles of Biological Circuits.

Amir, A., Kobiler, O., Rokney, A., Oppenheim, A.B., and Stavans, J. (2007). Noise in timing and precision of gene activities in a genetic cascade. Mol Syst Biol *3*, 71.

Andersen, J.B., Sternberg, C., Poulsen, L.K., Bjørn, S.P., Givskov, M., and Molin, S. (1998). New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Appl. Environ. Microbiol. *64*, 2240–2246.

Balaban, N.Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial Persistence as a Phenotypic Switch. Science (80-.). *305*, 1622 LP-1625.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709–1712.

Barrio, M., Burrage, K., Leier, A., and Tian, H. (2006). Oscillatory regulation of hes1: Discrete stochastic delay modelling and simulation. PLoS Comput. Biol. *2*, 1017–1030.

Basu, S., Mehreja, R., Thiberge, S., Chen, M.-T., and Weiss, R. (2004). Spatiotemporal control of gene expression with pulse-generating networks. Proc. Natl. Acad. Sci. U. S. A. *101*, 6355–6360.

Bennett, M.R., and Hasty, J. (2009). Microfluidic devices for measuring gene network dynamics in single cells. Nature *10*, 628–638.

Bernstein, J. a, Khodursky, A.B., Lin, P.-H., Lin-Chao, S., and Cohen, S.N. (2002). Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. Proc. Natl. Acad. Sci. U. S. A. *99*, 9697–9702.

Bilgin, N., Claesens, F., Pahverk, H., and Ehrenberg, M. (1992). Kinetic properties of Escherichia coli ribosomes with altered forms of S12. J. Mol. Biol. *224*, 1011–1027.

Brantl, S., and Wagner, E.G.H. (2002). An antisense RNA-mediated transcriptional attenuation mechanism functions in Escherichia coli. J. Bacteriol. *184*, 2740–2747.

Bratsun, D., Volfson, D., Tsimring, L.S., and Hasty, J. (2005). Delay-induced stochastic oscillations in gene regulation. Proc. Natl. Acad. Sci. U. S. A. *102*, 14593–14598.

Bremer, H., and Dennis, P.P. (1987a). Modulation of Chemical Composition and Other Parameters of the Cell by Growth Rate. Escherichia Coli Salmonella Cell. Mol. Biol. *2*, 1527–1542.

Bremer, H., and Dennis, P.P. (1987b). Modulation of Chemical Composition and Other Parameters of the Cell by Growth Rate. Escherichia Coli Salmonella Cell. Mol. Biol. *2*, 1527–1542.

Brewster, R.C., Jones, D.L., and Phillips, R. (2012). Tuning Promoter Strength through RNA Polymerase Binding Site Design in Escherichia coli. PLoS Comput. Biol. *8*.

Bustos, S. a, and Schleif, R.F. (1993). Functional domains of the AraC protein. Proc. Natl. Acad. Sci. U. S. A. *90*, 5638–5642.

Chabot, J.R., Pedraza, J.M., Luitel, P., and van Oudenaarden, A. (2007). Stochastic gene expression out-of-steady-state in the cyanobacterial circadian clock. Nature *450*, 1249–1252.

Chappell, J., Takahashi, M.K., and Lucks, J.B. (2015a). Creating small transcription activating RNAs. Nat. Chem. Biol. *11*, 1–9.

Chappell, J., Takahashi, M.K., and Lucks, J.B. (2015b). Creating small transcription activating RNAs. Nat. Chem. Biol. *11*, 1–9.

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S., et al. (2013a). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell *155*, 1479–1491.

Chen, Y.-J., Liu, P., Nielsen, A.A.K., Brophy, J.A.N., Clancy, K., Peterson, T., and Voigt, C.A. (2013b). Characterization of 582 natural and synthetic terminators and quantification of their design constraints. Nat. Methods *10*, 659–664.

Choi, P.J., Cai, L., Frieda, K., and Xie, X.S. (2008). A Stochastic Single-Molecule Event Triggers Phenotype Switching of a Bacterial Cell. Science (80-.). *322*, 442–446.

Cong, L., Ran, F., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P., Wu, X., Jiang, W., Marraffini, L., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. Science (80-.). *339*, 819–822.

Consortium, I.H.G.S. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Cookson, N. a, Mather, W.H., Danino, T., Mondragón-Palomino, O., Williams, R.J., Tsimring, L.S., and Hasty, J. (2011). Queueing up for enzymatic processing: correlated signaling through coupled degradation. Mol. Syst. Biol. *7*, 1–11.

Daniel, R., Rubens, J.R., Sarpeshkar, R., and Lu, T.K. (2013). Synthetic analog computation in living cells. Nature *497*, 619–623.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602–607.

Dicarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., and Church, G.M. (2013). Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. *41*, 4336–4343.

Diederich, L., Rasmussen, L.J., and Messer, W. (1992). New cloning vectors for integration into the lambda attachment site attB of the Escherichia coli chromosome. Plasmid *28*, 14–24.

Din, M.O., Danino, T., Prindle, A., Skalak, M., Selimkhanov, J., Allen, K., Julio, E., Atolia, E., Tsimring, L.S., Bhatia, S.N., et al. (2016). Synchronized cycles of bacterial lysis for in vivo delivery. Nature *536*, 81–85.

Dunlop, M.J., Cox, R.S., Levine, J.H., Murray, R.M., and Elowitz, M.B. (2008). Regulatory activity revealed by dynamic correlations in gene expression noise. Nat. Genet. *40*, 1493–1498.

Elf, J., Li, G.-W., and Xie, X.S. (2007). Probing transcription factor dynamics at the single-molecule level in a living cell. Science *316*, 1191–1194.

Elowitz, M.B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. Nature *403*, 335–338.

Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., Guptasarma, P., Spudich, J.L., McAdams, H.H., Heitzler, P., Ko, M.S., Fiering, S., et al. (2002). Stochastic gene expression in a single cell. Science *297*, 1183–1186.

Espah Borujeni, A., Channarasappa, A.S., and Salis, H.M. (2014). Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. Nucleic Acids Res. *42*, 2646–2659.

Falcon, C.M., and Matthews, K.S. (2000). Operator DNA sequence variation enhances high affinity binding by hinge helix mutants of lactose repressor protein. Biochemistry *39*, 11074–11083.

Fleishman, S.J., Whitehead, T.A., Ekiert, D.C., Dreyfus, C., Corn, J.E., Strauch, E.-M., Wilson, I. a, and Baker, D. (2011). Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. Science *332*, 816–821.

Friedland, A.E., Lu, T.K., Wang, X., Shi, D., Church, G., and Collins, J.J. (2009). Synthetic gene networks that count. Science *324*, 1199–1202.

Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch inEscherichia coli. Nature *403*, 339–342.

Gerdes, K., and Wagner, E.G.H. (2007). RNA antitoxins. Curr. Opin. Microbiol. *10*, 117–124.

Gerdes, K., Larsen, J.E.L., and Molin, S. (1985). Stable inheritance of plasmid R1 requires two different loci. J. Bacteriol. *161*, 292–298.

Ghusinga, K.R., and Singh, A. (2016). Optimal regulation of protein degradation to schedule cellular events with precision. In Proceedings of the American Control Conference, pp. 424–429.

Gilbert, L. a, Larson, M.H., Morsut, L., Liu, Z., Gloria, A., Torres, S.E., Stern-ginossar, N., Brandman, O., Whitehead, H., Doudna, J. a, et al. (2013). CRISPR-Mediated Modular RNA-Guided Regualtion of Transcription in Eukaryotes. Cell *154*, 442–451.

Gillespie, D.T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. *22*, 403–434.

Gillesple, D.T. (1977). Exact Stochastic Simulation of couple chemical reactions. J. Phys. Chem. *81*, 2340–2361.

Gordon, A., Colman-Lerner, A., Chin, T.E., Benjamin, K.R., Yu, R.C., and Brent, R. (2007). Single-cell quantification of molecules and rates using open-source microscope-based cytometry. Nat. Methods *4*, 175–181.

Green, A.A., Silver, P.A., Collins, J.J., and Yin, P. (2014). Toehold switches: De-novo-designed regulators of gene expression. Cell *159*, 925–939.

Gromiha, M.M., Thangakani, A.M., and Selvaraj, S. (2006). FOLD-RATE: Prediction of protein folding rates from amino acid sequence. Nucleic Acids Res. *34*.

Gupta, C., Lopez, J.M., Ott, W., Josic, K., and Bennett, M.R. (2013). Transcriptional delay stabilizes bistable gene networks. Phys. Rev. Lett. *111*.

Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., and Terns, M.P. (2009). RNA-Guided RNA Cleavage by a CRISPR RNA-Cas Protein Complex. Cell *139*, 945–956.

Hooshangi, S., Thiberge, S., and Weiss, R. (2005). Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. Proc. Natl. Acad. Sci. U. S. A. *102*, 3581–3586.

Hoynes-O'Connor, A., and Moon, T.S. (2016). Development of Design Rules for Reliable Antisense RNA Behavior in *E. coli*. ACS Synth. Biol. acssynbio.6b00036.

Hoynes-O'Connor, A., Hinman, K., Kirchner, L., and Moon, T.S. (2015). De novo design of heat-repressible RNA thermosensors in E. Coli. Nucleic Acids Res. *43*, 6166–6179.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262–1278.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol *31*, 227–229.

Jansen, R., Van Embden, J.D.A., Gaastra, W., and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. *43*, 1565–1575.

Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. a (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol *31*, 233–239.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptice Bacterial Immunity. Science *337*, 816–822.

Josić, K., López, J.M., Ott, W., Shiau, L., and Bennett, M.R. (2011). Stochastic delay accelerates signaling in gene networks. PLoS Comput. Biol. *7*.

Kampen, N.G. Van (2007). Stochastic Processes in Physics and Chemistry.

Kemmer, C., and Neubauer, P. (2006). Antisense RNA based down-regulation of RNaseE in E. coli. Microb. Cell Fact. *5*, 38.

Keung, A.J., Bashor, C.J., Kiriakov, S., Collins, J.J., and Khalil, A.S. (2014). Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. Cell *158*, 110–120.

Khlebnikov, A., Datsenko, K.A., Skaug, T., Wanner, B.L., and Keasling, J.D. (2001). Homogeneous expression of the PBAD promoter in Escherichia coli by constitutive expression of the low-affinity high-capacity araE transporter. Microbiology *147*, 3241–3247.

Kremers, G.J., Goedhart, J., Van Munster, E.B., and Gadella, T.W.J. (2006). Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET f??rster radius. Biochemistry *45*, 6570–6580.

Lee, Y.J., Hoynes-O'Connor, A., Leong, M.C., and Moon, T.S. (2016). Programmable control of bacterial gene expression with the combined CRISPR and antisense RNA system. Nucleic Acids Res. *44*, 2462–2473.

Levskaya, A., Chevalier, A.A., Tabor, J.J., Simpson, Z.B., Lavery, L.A., Levy, M., Davidson, E.A., Scouras, A., Ellington, A.D., Marcotte, E.M., et al. (2005). Synthetic biology: engineering Escherichia coli to see light. Nature *438*, 441–442.

Li, G.W. (2015). How do bacteria tune translation efficiency? Curr. Opin. Microbiol. *24*, 66–71.

Li, J., and Tooth, P. (1987). Size and shape of the Escherichia coli lactose permease measured in filamentous arrays. Biochemistry *26*, 4816–4823.

Li, M.Z., and Elledge, S.J. (2012). SLIC: A method for sequence- and ligation-independent cloning. Methods Mol. Biol. *852*, 51–59.

Li, G.-W., Berg, O.G., and Elf, J. (2009). Effects of macromolecular crowding and DNA looping on gene regulation kinetics. Nat. Phys. *5*, 294–297.

Li, G.W., Burkhardt, D., Gross, C., and Weissman, J.S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell *157*, 624–635.

Lucks, J.B., Qi, L., Mutalik, V.K., Wang, D., and Arkin, A.P. (2011). Versatile RNA-sensing transcriptional regulators for engineering genetic networks. Proc. Natl. Acad. Sci. U. S. A. *108*, 8617–8622.

Lutz, R., and Bujard, H. (1997). Independent and tight regulation of transcriptional units in escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Res. *25*, 1203–1210.

Marbach, A., and Bettenbrock, K. (2012). Lac operon induction in Escherichia coli: Systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. J. Biotechnol. *157*, 82–88.

Marraffini, L.A. (2015). CRISPR-Cas immunity in prokaryotes. Nature *526*, 55–61.

Mather, W., Bennett, M.R., Hasty, J., and Tsimring, L.S. (2009). Delay-induced degrade-and-fire oscillations in small genetic circuits. Phys. Rev. Lett. *102*.

Megerle, J.A., Fritz, G., Gerland, U., Jung, K., and Radler, J.O. (2008). Timing and Dynamics of Single Cell Gene Expression in the Arabinose Utilization System. Biophys. J. *95*, 2103–2115.

Mojica, F.J.M., Díez-Villaseñor, C., Soria, E., and Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol. Microbiol. *36*, 244–246.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. *60*, 174–182.

Moon, T.S., Lou, C., Tamsir, A., Stanton, B.C., and Voigt, C. a. (2012). Genetic programs constructed from layered logic gates in single cells. Nature *491*, 249–253.

Murray, A.W. (1992). Creative blocks: cell-cycle checkpoints and feedback controls. Nature *359*, 599–604.

Mutalik, V.K., Guimaraes, J.C., Cambray, G., Lam, C., Christoffersen, M.J., Mai, Q.-A., Tran, A.B., Paull, M., Keasling, J.D., Arkin, A.P., et al. (2013). Precise and reliable gene expression via standard transcription and translation initiation elements. Nat. Methods *10*, 354–360.

Na, D., Yoo, S.M., Chung, H., Park, H., Park, J.H., and Lee, S.Y. (2013). Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. Nat. Biotechnol. *31*, 170–174.

Nagai, T., Ibata, K., Park, E.S., Kubota, M., Mikoshiba, K., and Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. Nat. Biotechnol. *20*, 87–90.

Nagy, L., Gyetvai, G., and Nagy, G. (2009). Determination of the diffusion coefficient of monosaccharides with scanning electrochemical microscopy (SECM). Electroanalysis *21*, 542–549.

Narula, J., Kuchina, A., Lee, D.Y.D., Fujita, M., Süel, G.M., and Igoshin, O.A. (2015). Chromosomal Arrangement of Phosphorelay Genes Couples Sporulation and DNA Replication. Cell *162*, 328–337.

Nathan, L.P., Glenn, D.L., and Johan, V. (2016). Synchronous long-term oscillations in a synthetic gene circuit. Nature 1–4.

Nielsen, J., and Keasling, J.D. (2016). Engineering Cellular Metabolism. Cell *164*, 1185–1197.

Olson, E.J., Hartsough, L.A., Landry, B.P., Shroff, R., and Tabor, J.J. (2014). Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. Nat. Methods *11*, 449–455.

Orosz, G., Moehlis, J., and Murray, R.M. (2010). Controlling biological networks by time-delayed signals. Philos. Trans. A. Math. Phys. Eng. Sci. *368*, 439–454.

Ozbudak, E.M., Thattai, M., Lim, H.N., Shraiman, B.I., and Van Oudenaarden, A. (2004). Multistability in the lactose utilization network of Escherichia coli. Nature *427*, 737–740.

Paulsson, J. (2004). Summing up the noise in gene networks. Nature *427*, 415–418.

Pedraza, J.M., and van Oudenaarden, A. (2005). Noise propagation in gene networks. Science *307*, 1965–1969.

Peters, J.M., Colavin, A., Shi, H., Czarny, T.L., Larson, M.H., Wong, S., Hawkins, J.S., Lu, C.H.S., Koo, B.M., Marta, E., et al. (2016). A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. Cell *165*, 1493–1506.

Phillips, G. (1976). 3.5A Resolution Structure of L-Arabinose Binding Protein from E. coli. Rice University.

Pisponen, A., Mootse, H., Poikalainen, V., Kaart, T., Maran, U., and Karus, A. (2016). Effects of temperature and concentration on particle size in a lactose solution using dynamic light scattering analysis. Int. Dairy J. *61*, 205–210.

Prindle, A., Samayoa, P., Razinkov, I., Danino, T., Tsimring, L.S., and Hasty, J. (2012). A sensing array of radically coupled genetic “biopixels”. Nature *481*, 39–44.

Proshkin, S., Rahmouni, A.R., Mironov, A., and Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. Science *328*, 504–508.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173–1183.

Rathinam, M., Petzold, L.R., Cao, Y., and Gillespie, D.T. (2003). Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method. J. Chem. Phys. *119*, 12784–12794.

Ren, X., Sun, J., Housden, B.E., Hu, Y., Roesel, C., Lin, S., Liu, L., Yang, Z., Mao, D., Sun, L., et al. (2013). Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. Proc. Natl. Acad. Sci. U. S. A. *110*, 19012–19017.

Rosenfeld, N., and Alon, U. (2003). Response delays and the structure of transcription networks. J. Mol. Biol. *329*, 645–654.

Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S., and Elowitz, M.B. (2005). Gene regulation at the single-cell level. Science *307*, 1962–1965.

Shis, D.L., Hussain, F., Meinhardt, S., Swint-Kruse, L., and Bennett, M.R. (2014). Modular, Multi-Input Transcriptional Logic Gating with Orthogonal LacI/GalR Family Chimeras. ACS Synth. Biol. *3*, 645–651.

Siegele, D. a, and Hu, J.C. (1997). Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. Proc. Natl. Acad. Sci. U. S. A. *94*, 8168–8172.

Stanton, B.C., Nielsen, A. a K., Tamsir, A., Clancy, K., Peterson, T., and Voigt, C. a (2014). Genomic mining of prokaryotic repressors for orthogonal logic gates. Nat. Chem. Biol. *10*, 99–105.

Stricker, J., Cookson, S., Bennett, M.R., Mather, W.H., Tsimring, L.S., and Hasty, J. (2008). A fast, robust and tunable synthetic gene oscillator. Nature *456*, 516–519.

Strogatz, S.H. (1994). Nonlinear Dynamics and Chaos. Book 1–505.

Swint-Kruse, L., and Matthews, K.S. (2009). Allostery in the LacI/GalR family: variations on a theme. Curr. Opin. Microbiol. *12*, 129–137.

Tabor, J.J., Bayer, T.S., Simpson, Z.B., Levy, M., and Ellington, A.D. (2008). Engineering stochasticity in gene expression. Mol. Biosyst. *4*, 754–761.

Tabor, J.J., Salis, H.M., Simpson, Z.B., Chevalier, A.A., Levskaya, A., Marcotte, E.M., Voigt, C.A., and Ellington, A.D. (2009). A Synthetic Genetic Edge Detection Program. Cell *137*, 1272–1281.

Uphoff, S., Lord, N.D., Okumus, B., Potvin-Trottier, L., Sherratt, D.J., and Paulsson, J. (2016). Stochastic activation of a DNA damage response causes cell-to-cell mutation rate variation. Science (80-.). *351*, 1094–1097.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001). The sequence of the human genome. Science (80-.). *291*, 1304–1351.

Vogel, U., and Jensen, K.F. (1994a). The RNA chain elongation rate in Escherichia coli depends on the growth rate. J. Bacteriol. *176*, 2807–2813.

Vogel, U., and Jensen, K.F. (1994b). The RNA chain elongation rate in Escherichia coli depends on the growth rate. J. Bacteriol. *176*, 2807–2813.

Waaijers, S., Portegijs, V., Kerver, J., Lemmens, B.B.L.G., Tijsterman, M., van den Heuvel, S., and Boxem, M. (2013). CRISPR/Cas9-targeted mutagenesis in Caenorhabditis elegans. Genetics *195*, 1187–1191.

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/cas-mediated genome engineering. Cell *153*, 910–918.

Xiaodong, C. (2007). Exact stochastic simulation of coupled chemical reactions with delays. J. Chem. Phys. *126*.

Xie, K., and Yang, Y. (2013). RNA-Guided genome editing in plants using a CRISPR-Cas system. Mol. Plant *6*, 1975–1983.

Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R., and Benenson, Y. (2011). Multi-Input RNAi-Based Logic Circuit For Indentification of Specifc Cancer Cells. Science (80-.). *333*, 1307–1312.

Appendix

1. List of genetic parts

Part name	Type	DNA sequence
AraC	gene	Atggctgaagcgcaaaatgatcccctgctgccgggatactcgtttaacgcc catctggtggcgggtttaacgccgattgaggccaacggttatctcgatttttt atcgaccgaccgctgggaatgaaaggttatattctcaatctcaccattcgcgg tcaggggtggtgaaaaatcagggacgagaatttgtctgccgaccgggtg atattttgctgttcccgccaggagagattcatcactacggtcgtcatccggag gctcgcgaatggtatcaccagtgggtttactttcgtccgcgcgcctactggca tgaatggcttaactggccgtcaatatttgccaatacgggtttctttcgcccgga tgaagcgcaccagccgcatttcagcgacctgtttgggcaaatcattaacgcc gggcaaggggaagggcgctattcggagctgctggcgataaatctgcttga gcaattgttactgcggcgcatggaagcgattaacgagtcgctccatccaccg atggataatcgggtacgcgaggcttgtcagtacatcagcgatcacctggca gacagcaattttgatatcgccagcgtcgcacagcatgtttgcttgtcgccgtc gcgtctgtcacatcttttccgccagcagttagggattagcgtcttaagctggc gcgaggaccaacgcattagtcaggcgaagctgcttttgagcactacccgga tgcctatcgccaccgtcggtcgcaatgttggttttgacgatcaactctatttctc gcgagtatttaaaaaatgcaccggggccagcccgagcgagtttcgtgccg gttgtgaagaaaaagtgaatgatgtagccgtcaagttgtcatga
LacI	gene	Atgaaaccagtaacgttatacgatgtcgcagagtatgccggtgtctcttatca gaccgtttcccgcgtggtgaaccaggccagccacgtttctgcgaaaacgcg ggaaaaagtggaagcggcgatggcggagctgaattacattcccaaccgcg tggcacaacaactggcgggcaaacagtcgttgctgattggcgttgccacct ccagtctggccctgcacgcgccgtcgcaaattgtcgcggcgattaaatctcg cgccgatcaactgggtgccagcgtggtggtgtcgatggtagaacgaagcg gcgtcgaagcctgtaaagcggcggtgcacaatcttctcgcgcaacgcgtca gtgggctgatcattaactatccgctggatgaccaggatgccattgctgtgga agctgcctgcactaatgttccggcgttatttcttgatgtctctgaccagacacc catcaacagtattattttctcccatgaagacggtacgcgactgggcgtggag catctggtcgcattgggtcaccagcaaatcgcgctgttagcgggcccattaa gttctgtctcggcgcgtctgcgtctggctggctggcataaatatctcactcgc aatcaaattcagccgatagcggaacgggaaggcgactggagtgccatgtc

		cggttttcaacaaaccatgcaaatgctgaatgagggcatcgttcccactgcg atgctggttgccaacgatcagatggcgctgggcgcaatgcgcgccattacc gagtccgggctgcgcgttggtgcggatatctcggtagtgggatacgacgat accgaagacagctcatgttatatcccgccgttaaccaccatcaaacaggattt tcgcctgctggggcaaaccagcgtggaccgcttgctgcaactctctcaggg ccaggcggtgaagggcaatcagctgttgcccgtctcactggtgaaaagaaa aaccaccctggcgcccaatacgcaaaccgcctctccccgcgcgttggccg attcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagt ga
sfYFP	gene	Atgcgtaaaggcgaagagctgttcactggtgtcgtccctattctggtggaac tggatggtgatgtcaacggtcataagttttccgtgcgtggcgagggtgaagg tgacgcaactaatggtaaactgacgctgaagttcatctgtactactggtaaact gccggtaccttggccgactctggtaacgacgctgacttatggtgttcagtgct ttgctcgttatccggaccatatgaagcagcatgacttcttcaagtccgccatgc cggaaggctatgtgcaggaacgcacgatttcctttaaggatgacggcacgt acaaaacgcgtgcggaagtgaaatttgaaggcgataccctggtaaaccgca ttgagctgaaaggcattgactttaaagaagacggcaatatcctgggccataa gctggaatacaatttaacagccacaatgtttacatcaccgccgataaacaaa aaaatggcattaaagcgaattttaaaattcgccacaacgtggaggatggcag cgtgcagctggctgatcactaccagcaaaacactccaatcggtgatggtcct gttctgctgccagacaatcactatctgagcTACcaaagcgttctgtctaaa gatccgaacgagaaacgcgatcatatggttctgctggagttcgtaaccgca gcgggcatcacgcatggtatggatgaactgtacaaatga
TetR	gene	Atggctggttctcgcagaaagaaacatatccatgaaatcccgccccgaattc atatgtctagattagataaaagtaaagtgattaacagcgcattagagctgctta atgaggtcggaatcgaaggtttaacaacccgtaaactcgcccagaagctag gtgtagagcagcctacattgtattggcatgtaaaaaataagcgggctttgctc gacgccttagccattgagatgttagataggcaccatactcacttttgccctta gaaggggaaagctggcaagatttttacgtaataacgctaaaagttttagatg tgctttactaagtcatcgcgatggagcaaaagtacatttaggtacacggccta cagaaaaacagtatgaaactctcgaaaatcaattagccttttatgccaacaa ggtttttcactagagaatgcattatatgcactcagcgctgtggggcattttactt taggttgcgtattggaagatcaagagcatcaagtcgctaaagaagaaaggg aaacacctactactgatagtatgccgccattattacgacaagctatcgaattat ttgatcaccaaggtgcagagccagccttcttattcggccttgaattgatcatat gcggattagaaaaacaacttaaatgtgaaagtgggtcttaa
$P_{A1lacO-1}$	promoter	Aaatttatcaaaaagagtgttgacttgtgagcggataacaatgatacttagatt caattgtgagcggataacaatttcaca

P_{BAD}	promoter	Acattgattatttgcacggcgtcacactttgctatgccatagcatttttatccataagattagcggatcctacctgacgctttttatcgcaactctctactgtttctcc
$P_{lac/ara}$	promoter	Tgtgtggaattgtgagcggataacaatttcacacagggccctcggacaccgaggagaatgtcaagaggcgaacacacaacgtcttggagcgccagaggaggaacgagctaaaacggagctttttgccctgcgtgaccagatcccggagttggaaaacaatgaaaaggcccccaaggtagttatccttaaaaaagccacagcatacatcctgtccgtccaagcagaggagcaaaagctcatttctgaagaggacttgttgcggaaacgacgagaacagttgaaacacaaacttgaacagctacggaactcttgtgcgtaaggaaaagtaaggaaaacgattccttctaacagaaatgtcctgagcaatcacctatgaactgtcgactcgagcatagcatttttatccataagattagcggatcctaagctttacaattgtgagcgctcacaattatgatagattcaattgtgagcggataacaattgcatgc
P_{Tet}	promoter	Tccctatcagtgatagagattgacatccctatcagtgatagagatactgagcac
$P_{LlacO-1}$	promoter	Aattgtgagcggataacaattgacattgtgagcggataacaagatactgagcac
J23100	promoter	Ttgacggctagctcagtcctaggtacagtgctagc
BCD	RBS	Cacttaaaaaggagatcaacaatgaaagcaattttcgtactgaaacatcttaatcatgctaaggagaaatactagt
B0034	RBS	Aaagaggagaaa
LAA tag	peptide	Actagtgcagcgaacgacgaaaattacgcccttgcagcg
Antisense YFP	RNA	Taagttttccgtgctcaccttcaccctcgccacgcacggaaaacttatgaccgttgacatcaccatccagttcc
Antisense mCherry	RNA	Agttcatgcgcttcccctccatgtgcaccttgaagcgcatgaactccttgatgatggccatgttatcctcctcg
Antisense AraC	RNA	Ctcgtttaacgccctaaacccgccaccagatgggcgttaaacgagtatcccggcagcaggggatcattttgcgc

Antisense LacI	RNA	Ttctgcgaaaacgcccatcgccgcttccactttttcccgcgttttcgcagaaacgtggctggcctggttcaccacgcg