Submitted, 2015 Winter q-bio Conference (5 Nov 2014)
http://www.cds.caltech.edu/~murray/papers/ghm15-wqbio.html

Systematic Design and Implementation of a Novel Synthetic Fold-Change Detector Biocircuit *In Vivo*

Shaobin Guo, Yutaka Hori, Richard Murray

Nov 5, 2014

Biological signaling systems not only detect the absolute levels of the signals, but are also able to sense the fold-changes of the signals. The ability to detect fold-changes provides a powerful tool for biological organisms to adapt to the changes in environment. Here we present the first novel synthetic fold-change detector (FCD) circuit built from ground up *in vivo*. We systematically designed the FCD circuit *in silico*, prototyped it in cell-free transcription-translation platform (TX-TL), and eventually implemented it in *E. coli* cells. We were able to show that the FCD circuit can not only generate pulse-like behavior in response to input, but also produce the same pulse response with inputs of the same fold-change, despite of different absolute signal levels.

1 Introduction

In common sense, biological sensory systems detect the absolute signal levels of the inputs and produce corresponding outputs. However, Weber's Law in sensory physiology shined some creative light on how some biological gene regulation networks may actually work in cells [1]. In Wnt pathway, fold-change in the level of β-catenin induced by Wnt, and not the final level itself, was shown to be likely the actual reporter of Wnt signal that is sensed by the downstream transcriptional system [2]. In ERK signaling system, upon ligand stimulation, cells showed a fold-change response, where peak nuclear accumulation of ERK2 is proportional to basal level in each cell [3]. These evidences suggest that some gene regulation networks are using fold-change detection for the virtue of robust response to real signals among background noise or variations.

Gene regulation networks are composed of a small set of recurring interaction patterns called network motifs [4, 5]. Among all the common gene regulation network motifs, incoherent type-1 feedforward loop (FFL) is one of the widely used and interesting ones [6]. The FFL is composed of two input transcription factors, one (x) of which activates the other (y), both jointly regulating (activating or repressing) a target gene (z) (Figure 1A). Because there is a time delay between the activation of z by x and the repression of z by y, it has been shown that this type of FFL can generate a temporal pulse of z response. Furthermore, previous theoretical studies have demonstrated that the FFL could have possessed another feature, which is generating a response to fold-changes in the input signal, rather than absolute levels [7]. Experiments supporting the theory have been performed *in vitro* using genelet system, where they demonstrated the incoherent type-1 FFL is able to show exact adaptation and fold-change detection [8].

Here we demonstrate that by systematically designing and building an incoherent type-1 FFL, we are able to show fold-change detection with this circuit. First we showed a simple mathematical model for the design and simulation results indicating fold-change detection. Next we prototyped and tested the FFL circuit in cell-free transcription-translation breadboard platform [9, 10]. Finally we implemented the FFL *in vivo* and demonstrated that instead of sensing the absolute level of the inputs, it actually detects the fold-change between the previous and current inputs. This first successful application of a novel synthetic fold-change detector *in vivo* will help the understanding of more complex fold-change detectors naturally existing in cells.

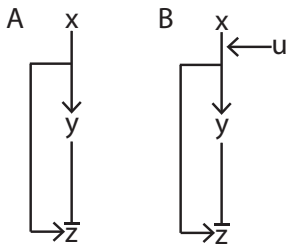

Figure 1: Feedforward loop diagram

2 Mathematical Model

Before we actually built and tested the fold-change detector, we would like to first test the feasibility of the idea *in silico*. Previous studies have extensively demonstrated how to get fold-change response from a incoherent type-1 feedforward loop mathematically [7, 8]. So we focus on the differences here. In reality, we would like to have control over the input x (Figure 1A). We only want x to turn on y and z when we ask it to. But to

keep things simpler, it would be good that x can be seen as a constant, meaning when the activation starts, the concentration of x has already reached steady state and x is in abundance. Then we would need another factor (u), which we have fully control of, to turn the activation on (Figure 1B). In this case, the activation of y and z is positively correlated with u and u can be seen as an inducer that is required by x to activate downstream genes. And we can write down the following ODEs:

$$\dot{y} = -d_1 y + \frac{c_1 u}{K_1 + u} \tag{1a}$$

$$\dot{z} = -d_2 z + \frac{c_2 u}{(K_1 + u)(\frac{y}{K_2} + 1)} \tag{1b}$$

in which d_1 and d_2 are the degradation plus dilution rates of y and z, respectively. And they are essentially the same when tagged with the same degradation tag. c_1 and c_2 are the production rates of y and z, which include transcription, translation, folding rates and so on. K_1 is the binding rate constant of $[x : u]$ complex to y and z, which should be the same for both y and z if they share the same activation mechanism. K_2 is the repression rate constant of y on z.

To achieve fold-change detection, we needed to prove that z is only determined by the fold-change between u_0 and u, where u_0 is the current inducer concentration and u is the next inducer concentration. First we needed u in its linear induction range and then the equations can be simplified as shown below:

$$\dot{y} = -d_1 y + \frac{c_1 u}{K_1} \tag{2a}$$

$$\dot{z} = -d_2 z + \frac{c_2 u}{K_1(\frac{y}{K_2} + 1)} \tag{2b}$$

Secondly, we wanted y to be a strong repressor on z, meaning y is significantly larger than K_2 and then we can get the following $\dot{z}$ equation:

$$\dot{z} = -d_2 z + \frac{c_2 K_2 u}{K_1 y} \tag{3}$$

Let $c_1/K_1 = \beta_1$, $d_1 = \alpha_1$, $c_2 K_2/K_1 = \beta_2$ and $d_2 = \alpha_2$, the equations become:

$$\dot{y} = -\alpha_1 y + \beta_1 u \tag{4a}$$

$$\dot{z} = -\alpha_2 z + \beta_2 \frac{u}{y} \tag{4b}$$

where equations 4a and 4b are exactly the same as those demonstrated in [7]. After performing the same dimensionless variable replacement, we would get

$$\frac{dY}{d\tau} = F - Y \tag{5a}$$

$$r\frac{dZ}{d\tau} = \frac{F}{Y} - Z \tag{5b}$$

where $F = u/u_0$, $\tau = \alpha_1 t$, $r = \alpha_1/\alpha_2$, $Y = \frac{y}{\beta_1 u_0/\alpha_1}$ and $Z = \frac{z}{\beta_2\alpha_1\alpha_2/\beta_1}$. So we could have the dynamics of Z only depends on F, which is the fold-change of u.

Next step we need to estimate the parameters for equations 1a and 1b. Degradation and dilution terms d_1 and d_2 can be estimated as $\frac{ln2}{t_{1/2}}$ where $t_{1/2}$ is protein half-life (about 1 hour) [11]. Then we chose physiologically reasonable parameters for the other parameters with constraints such as K_2 should be significantly smaller than y. While parameter sweeping, we found that c_2 needs to much larger than c_1 (such as 100 fold) to get observable output z as the repression from y is very strong (K_2 small). With a set of rational parameters, we were able to see fold change detection in simulation (Figure 2).

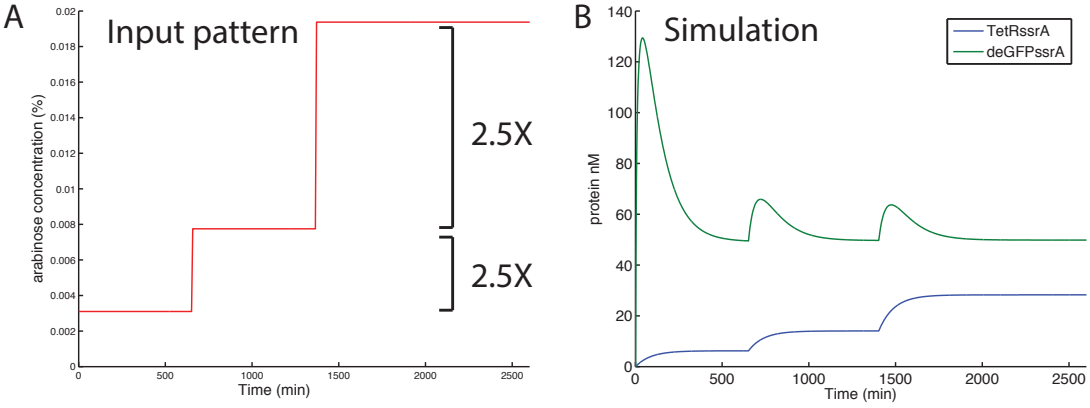

Figure 2: FCD simulation. **A**: Fold-change input pattern with two 2.5 fold-change. **B**: Simulation results of FCD output pattern. $u = 0.0031, d_1 = d_2 = ln2/60, c_1 = 0.001, c_2 = 500c_1, K_1 = 0.0004$ and $K_2 = 0.0001$.

3 Build a novel synthetic FFL circuit *in vitro* and *in vivo*

After having a working model, we designed and built the incoherent type-1 FFL shown in Figure 3A. The FFL is composed of two input transcription factors, one of which activates the other, both jointly regulating (activating or repressing) a target gene [12].

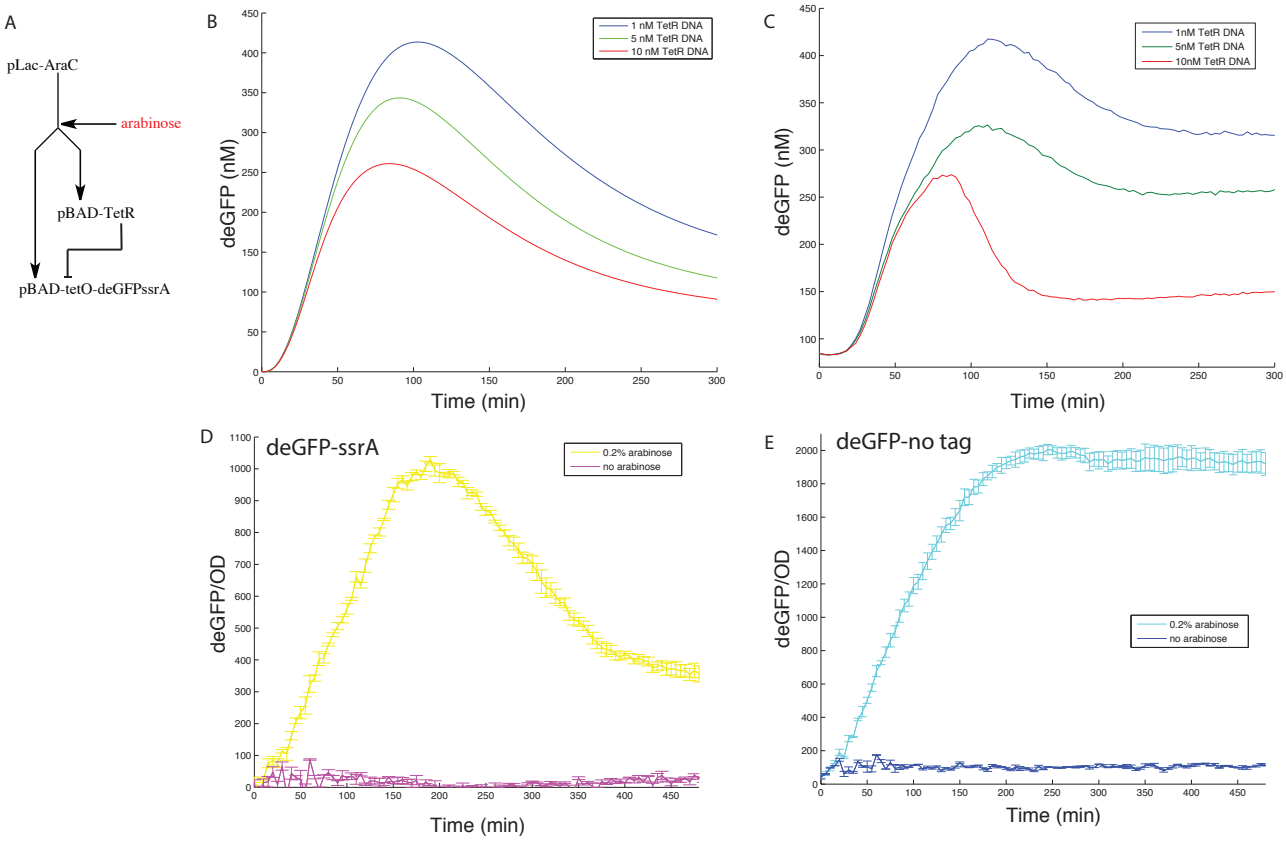

Figure 3: FFL circuit design and prototyping. **A**: FFL circuit diagram. **B**: Simulation results of FFL in TX-TL toolbox. **C**: FFL experiment in TX-TL (linear DNA): 10nM AraC linear DNA, 10nM deGFP linear DNA, 10nM ClpX linear DNA, 0.2% arabinose and varied TetR linear DNA concentrations. **D,E**: FFL experiment *in vivo*: KL740 cells transformed with AraC-TetR-deGFP(ssrA) plasmids and grown in defined MOPS media. **D**: deGFP is tagged with ssrA. **E**: deGFP does not have a tag. GFP signals were normalized by OD600 to get the fluorescence reading for each cell. Inducer arabinose concentration is none or 0.2%. Experiment ran at 37 °C on BIOTEK Synergy H1 Hybrid Multi-Mode Microplate Reader.

In cell strains which don't overexpress LacI protein, pLac-AraC will continuously ex-

press transcription factor AraC, which binds to promoter $pBAD$ and activates the transcription of downstream genes (TetR and deGFP) only in the presence of arabinose. Also as mentioned in the model, the AraC concentration can be seen as a constant and abundant. On the other hand, transcription factor TetR binds to operator site *tet* and represses the transcription of deGFP. Because of the time delay between the activation by AraC-arabinose and the repression by TetR, deGFP gene will first be transcribed and after TetR proteins accumulate to certain level, the transcription ceases. At the same time, the ssrA tag on deGFP protein leads to degradation of the protein. As a result, deGFP fluorescence signal, which can be measured and quantified, will increase in the beginning and then decrease, generating a pulse (simulated in Figure 3B) [13].

We first designed and ordered primers (a day before) to amplify the coding sequences for AraC, TetR and deGFPssrA. Then we used GoldenBraid assembly method to stitch specific promoters, ribosome binding sites (RBSs), coding sequences (CDSs) and terminators together with plasmid vectors [14]. After 1 hour incubation, we amplified the linear DNAs containing P-R-C-T-V via PCR reactions. Then we used these linear DNAs to run experiments in TX-TL. From start to finish, one experiment cycle will be less than a day. Figure 3C showed the experimental results from TX-TL experiments. All the curves were consistent with the simulation: GFP signal first increased as a result of AraC-arabinose activation; then after TetR proteins accumulated to certain amount, they repressed the transcription of deGFPssrA and at the same time, ClpX protein, which is an ATPase, unfolded tagged deGFP proteins and brought the GFP signal down. As we can see in the figure, the more TetR DNAs were added, the lower GFP signal was and the faster peak appeared. This is because increasing TetR DNA concentrations will reduce the time needed to produce enough TetR proteins to shut down the GFP transcription. However, it is easy to notice that the GFP signals at steady state for different TetR DNA concentrations were different. This is because TX-TL reaction has limited resource, including RNA polymerase, NTPs, ribosomes and amino acids [15]. The energy required for ClpX to unfold tagged deGFP proteins will run out gradually and the more DNAs are added, the faster the resource gets used up. In consequence, they showed different steady state signals.

After we tested linear DNA version of the FFL in TX-TL and found a working design, we assembled these linear DNAs into one plasmid using GoldenBraid assembly method in order to implement the circuit *in vivo*. we transformed the plasmid into *E. coli* cells KL740 [16]. Besides, to make sure the reduction of GFP signal *in vivo* is a result of GFP degradation instead of the dilution introduced by cell division, we had a control circuit in which deGFP is not tagged with ssrA. Figure 3D and E showed the results from *in vivo* experiments of the FFL (Figure 3D) and the control circuit (Figure 3E). The dynamics shown in Figure 3D were clearly consistent with those shown in TX-TL experiments, meaning the FFL prototyped in TX-TL indeed showed the same behavior *in vivo*. In contrast, the control circuit in Figure 3E didn't have a pulse-like behavior, suggesting that the pulse we saw was not a result of dilution but was from the degradation of GFP proteins.

4 Implement the novel FFL as FCD *in vivo*

Getting a novel working FFL is just the first step. Based on this structure, we then modified and optimized the FFL to turn it into a fold-change detector (FCD) circuit. Based on our model, TetR should have the same degradation rate constant as deGFP. So we put the same ssrA tag on TetR (Figure 4A). The result is that at steady state, TetR concentration should be constant. The single plasmid containing all three components was transformed into MG1655 [17]. Also model suggested c_2 to be much larger than c_1, we used a weak ribosome binding site and very strong ribosome binding site (more than 200 fold difference between them) [18]. Cells were grown at 29 °C to make the doubling time longer and extend the time window for circuit dynamics. And for the same reason, cells were grown in M9 minimal defined media to ensure longer experimental time [19]. At $time = 0$, cells were induced with different concentrations of inducer arabinose (from 0% to 1.9%).

First we only did one induction just to see the dynamic range of the pulses as a function of arabinose concentration. As shown in Figure 4B, as arabinose concentration increased, the pulse got higher and higher. But after 0.0488%, the pulse heights, which were extracted from Figure 4B and plotted in Figure 4C, started to decrease and the steady state value of the curve became larger, suggesting the system has hit the arabinose concentration limitation. Adding more arabinose won't necessarily increase the peak.

After figuring out the range of effective arabinose concentration (0.0002% to 0.0488%), we tested the circuit's fold-change detection capability. We first grew cells overnight in 10mL round-bottom culture tubes without any inducers, and then diluted the cells into a 48-wells plate to around $OD600 = 0.01$. For each column, there were eight different conditions, where the arabinose concentration of the initial induction was different: from 0.0002% to 0.0488% in 2.5 fold change increment). For each row there were three different experimental conditions and each condition has two repeats. The three conditions were: 1. one induction with the initial arabinose concentration; 2. two inductions with the second arabinose concentration being 2.5 fold of the initial one, and 3. three inductions with the third arabinose concentration being 6.25 fold of the initial one. The second and third inductions were introduced after the pulses reached steady state at basal level.

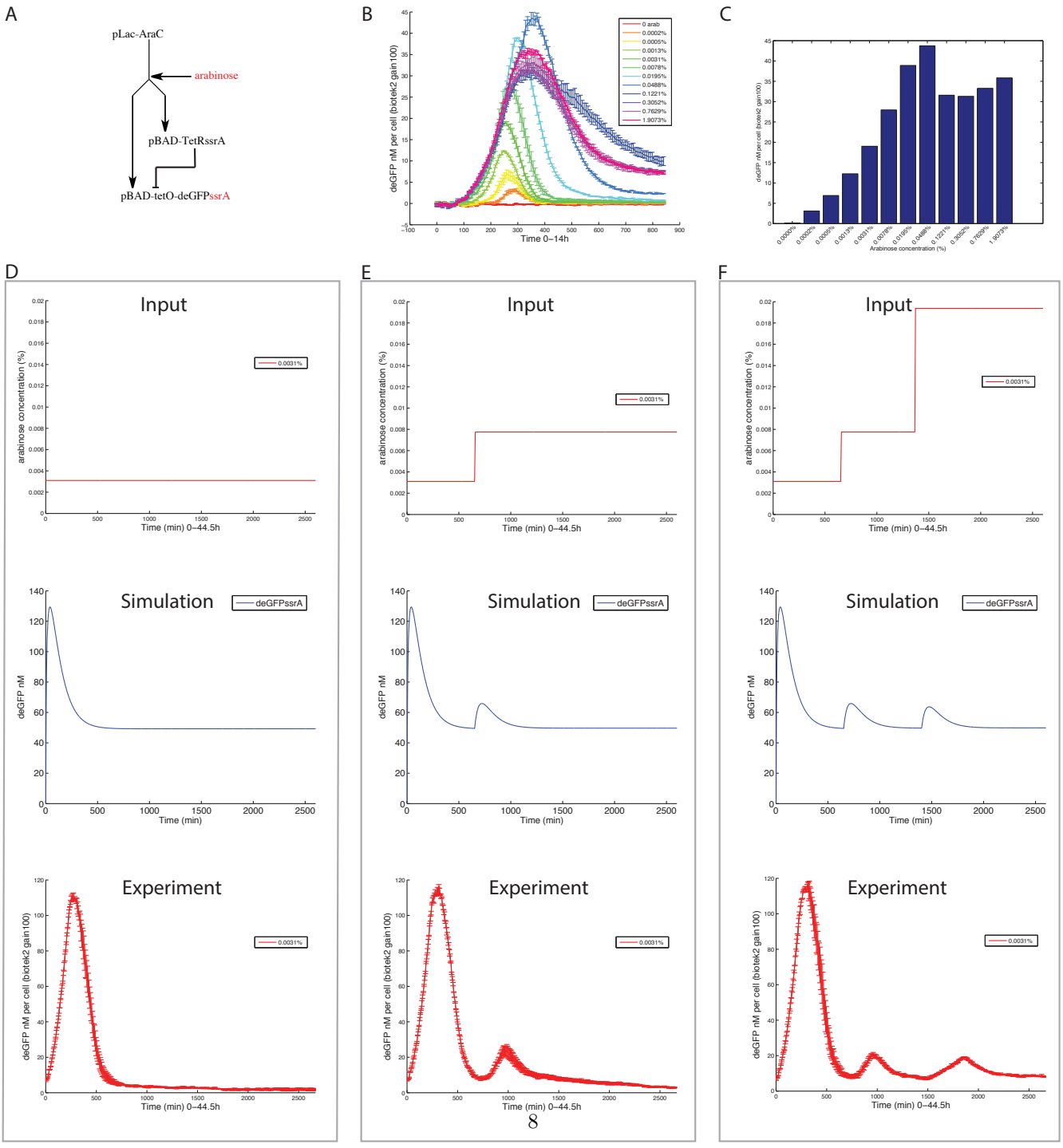

Figure 4: FCD circuit design, simulation and implementation *in vivo*. **A**: FCD circuit diagram. Continued on next page.

Figure 4: **B**: FCD responded to different initial inducer concentrations. **C**: Peak heights extracted from **B** were plotted against the initial inducer concentrations. **D-F**: Fold-change detection experiment in *E. coli* MG1655 with an initial inducer arabinose concentration of 0.0031% or $0.2mM$. Three input patterns were tested. Next input step was executed after the previous pulse reached steady state. Ideal output patterns were shown in corresponding simulation results. Both deGFP and TetR concentrations were plotted for simulation. Experiment ran at 29 °C continuously, except for adding more inducers, on BIOTEK Synergy H1 Hybrid Multi-Mode Microplate Reader.

As shown in Figure 4D, the input pattern was just a constant after $time = 0$. The simulation showed that a pulse was generated in response to the input (FFL response). The *in vivo* experiment suggested the same result. In Figure 4E, the input pattern was a two-step function where the second step was 2.5 fold higher than the first step. Simulation result showed that after the initial pulse, there was another pulse generated as a result of the second induction. The peak of the second pulse was significant lower compared to the first one. This is because the first pulse was in theory an infinite fold-change and the second pulse was a 2.5 fold-change. The experimental result shown below matched qualitatively with the simulation. But to verify the circuit can do fold-change detection, we needed one more step where the third step was 2.5 fold higher than the second step. If both pulses looked the same, it would mean the circuit was detecting fold-change instead of the absolute arabinose concentration change.

As shown in Figure 4F, the input pattern was a three-step function as we mentioned above. The simulation result suggested that as the second and third inductions had the same fold-change, the peaks should be similar. And indeed the circuit detected fold-change as shown in the experimental result, where the last two pulses were almost the same. These two pulses had very different absolute arabinose concentrations; but instead of showing peaks with different heights, they ended up with the same height. This was strong evidence that our FFL is actually a FCD circuit.

5 Conclusion

Fold-change detection is naturally existing in some biological sensing systems. Utilizing this specific function of biological sensors is still in progress. To fully understand how a thing work, we need to know how to build it. We showed a working novel synthetic fold-change detector built from scratch. By testing the idea *in silico*, *in vitro* and eventually *in vivo*, we successfully demonstrated how a circuit can respond to the fold-change in inputs instead of the absolute level changes in inputs. Inputs with the same fold-change will generate the same pulse-like behavior, despite of their differences in absolute values.

Having fold-change detection will enable organisms to filter out background noise and focus on the real signal changes. The ability of adapting to environmental changes is also beneficial to organisms, giving them the capability to survive in changing environments. Our work on modifying simple synthetic biocircuits to do more complicated things shines lights on the idea that bigger isn't always better (bigger brain is desirable but what if 90% of the current brain is still in idle). FCD might also be used as the upstream detector for some more complicated circuits, such as fold-change threshold detector, where only when the fold-change is large enough, the downstream reporter will be turned on. We are looking forward to more possibilities with this novel synthetic biocircuit module.

References

[1] E Weber, "Der tatsinn und das gemeingefuhl", *Handwörterbuch der Physiologie. Wilhelm Engelmann, Leipzig*, 1846.

[2] Lea Goentoro and Marc W Kirschner, "Evidence that fold-change, and not absolute level, of β-catenin dictates wnt signaling", *Molecular cell*, vol. 36, no. 5, pp. 872–884, 2009.

[3] Cellina Cohen-Saidon, Ariel A Cohen, Alex Sigal, Yuvalal Liron, and Uri Alon, "Dynamics and variability of erk2 response to egf in individual living cells", *Molecular cell*, vol. 36, no. 5, pp. 885–893, 2009.

[4] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon, "Network motifs: simple building blocks of complex networks", *Science*, vol. 298, no. 5594, pp. 824–827, 2002.

[5] Shai S Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon, "Network motifs in the transcriptional regulation network of escherichia coli", *Nature genetics*, vol. 31, no. 1, pp. 64–68, 2002.

[6] Patrick Eichenberger, Masaya Fujita, Shane T Jensen, Erin M Conlon, David Z Rudner, Stephanie T Wang, Caitlin Ferguson, Koki Haga, Tsutomu Sato, Jun S Liu, et al., "The program of gene transcription for a single differentiating cell type during sporulation in bacillus subtilis", *PLoS biology*, vol. 2, no. 10, pp. e328, 2004.

[7] Lea Goentoro, Oren Shoval, Marc W Kirschner, and Uri Alon, "The incoherent feedforward loop can provide fold-change detection in gene regulation", *Molecular cell*, vol. 36, no. 5, pp. 894–899, 2009.

[8] Jongmin Kim, Ishan Khetarpal, Shaunak Sen, and Richard M Murray, "Synthetic circuit for exact adaptation and fold-change detection", *Nucleic acids research*, p. gku233, 2014.

[9] Zachary Z Sun, Clarmyra A Hayes, Jonghyeon Shin, Filippo Caschera, Richard M Murray, and Vincent Noireaux, "Protocols for implementing an escherichia coli based tx-tl cell-free expression system for synthetic biology", *JoVE (Journal of Visualized Experiments)*, , no. 79, pp. e50762–e50762, 2013.

[10] Zachary Z Sun, Enoch Yeung, Clarmyra A Hayes, Vincent Noireaux, and Richard M Murray, "Linear dna for rapid prototyping of synthetic biological circuits in an escherichia coli based tx-tl cell-free system", *ACS synthetic biology*, 2013.

[11] Christian Trötschel, Stefan P Albaum, and Ansgar Poetsch, "Proteome turnover in bacteria: current status for corynebacterium glutamicum and related bacteria", *Microbial biotechnology*, vol. 6, no. 6, pp. 708–719, 2013.

[12] Shmoolik Mangan and Uri Alon, "Structure and function of the feed-forward loop network motif", *Proceedings of the National Academy of Sciences*, vol. 100, no. 21, pp. 11980–11985, 2003.

[13] ZA Tusa, Vipul Singhal, Jongmin Kim, and Richard M Murray, "An in silico modeling toolbox for rapid prototyping of circuits in a biomolecular ?breadboard? system", in *2013 Conference on Decision and Control*. IEEE Florence, Italy, 2013.

[14] Alejandro Sarrion-Perdigones, Erica Elvira Falconi, Sara I Zandalinas, Paloma Juárez, Asun Fernández-del Carmen, Antonio Granell, and Diego Orzaez, "Golden-braid: an iterative cloning system for standardized assembly of reusable genetic modules", *PloS one*, vol. 6, no. 7, pp. e21622, 2011.

[15] Dan Siegal-Gaskins, Vincent Noireaux, and Richard M Murray, "Biomolecular resource utilization in elementary cell-free gene circuits", in *American Control Conference (ACC), 2013*. IEEE, 2013, pp. 1531–1536.

[16] David J Weber, Apostolos G Gittis, Gregory P Mullen, Chitrananda Abeygunawardana, Eaton E Lattman, and Albert S Mildvan, "Nmr docking of a substrate into the x-ray structure of staphylococcal nuclease", *Proteins: Structure, Function, and Bioinformatics*, vol. 13, no. 4, pp. 275–287, 1992.

[17] JS Edwards and BO Palsson, "The escherichia coli mg1655 in silico metabolic genotype: its definition, characteristics, and capabilities", *Proceedings of the National Academy of Sciences*, vol. 97, no. 10, pp. 5528–5533, 2000.

[18] Vivek K Mutalik, Joao C Guimaraes, Guillaume Cambray, Colin Lam, Marc Juul Christoffersen, Quynh-Anh Mai, Andrew B Tran, Morgan Paull, Jay D Keasling, Adam P Arkin, et al., "Precise and reliable gene expression via standard transcription and translation initiation elements", *Nature methods*, vol. 10, no. 4, pp. 354–360, 2013.

[19] Joseph Sambrook, David W Russell, and David W Russell, *Molecular cloning: a laboratory manual (3-volume set)*, vol. 999, Cold spring harbor laboratory press Cold Spring Harbor, New York:, 2001.