The timing of transcriptional regulation in synthetic gene circuits

Yu-Yu Cheng[1], Andrew J. Hirning[1], Krešimir Josić[1,2,3], and Matthew R. Bennett[1,4,†]

[1]Department of Biosciences, Rice University, Houston, TX 77005, USA.
[2]Department of Mathematics, University of Houston, Houston, TX 77204, USA.
[3]Department of Biosciences, University of Houston, Houston, TX 77204, USA.
[4]Department of Bioengineering, Rice University, Houston, TX 77005, USA.
[†]Corresponding author. E-mail: matthew.bennett@rice.edu

Abstract

Transcription factors and their target promoters are central to synthetic biology. By arranging these components into novel gene regulatory circuits, synthetic biologists have been able to create a wide variety of phenotypes, including bistable switches, oscillators, and logic gates. However, transcription factors (TFs) do not instantaneously regulate downstream targets. After the gene encoding a TF is turned on, the gene 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 "signaling time," is a critical aspect in the design of dynamic regulatory networks, yet it remains poorly characterized. In this work, we measured the signaling time of two TFs in *Escherichia coli* commonly used in synthetic biology: the activator AraC and the repressor LacI. We found that signaling times can range from a few to tens of minutes, and are affected by the expression rate of the TF. Our single-cell data also show that the variability of the signaling time increases with its mean. To validate these signaling time measurements, we constructed a two-step genetic cascade, and showed that the signaling time of the full cascade can be predicted from those of its constituent steps. These results provide concrete estimates for the timescales of transcriptional regulation in living cells, which are important for understanding the dynamics of synthetic transcriptional gene circuits.

Introduction

One aspect of synthetic biology is the rearrangement of regulatory mechanisms within cells to elicit novel phenotypes. While post-transcriptional regulatory mechanisms are sometimes used[1,2], the majority of synthetic gene circuits in *E. coli* primarily rely on transcription factors (TFs) and their target promoters. By engineering novel transcriptional regulatory topologies, synthetic biologists have created a vast array of genetic circuits[3], including toggle switches[4,5], oscillators[6,7], logic gates[8,9], dose-response linearizers[10], and multicellular systems[11-14]. Yet, despite the increasing scale and complexity of synthetic transcriptional circuits[15] and a deep theoretical understanding of the consequences of finite signaling times[16], a precise experimental understanding of the dynamics of transcriptional signaling is lacking.

For many types of gene circuits, transcriptional signaling times[17] do not play an important role. If a circuit is designed to only operate at steady state then transcriptional signaling times only add to the transient time scales generally governed by cell growth and proteolysis[18]. For instance, dose-response linearizers[10] simplify the relationship between the input concentration of inducer (the dose) and the output steady-state concentration of protein (the response). Provided the time it takes to reach steady state protein levels is inconsequential, transcriptional signaling times do not need to be taken into account during the design process.

If the dynamics of a synthetic gene circuit is important to its function, transcriptional signaling times can drastically alter their behavior. This is especially true for many (though not all) gene circuits that have transcriptional feedback loops. The oscillations in some synthetic circuits, for example, are thought to result from transcriptional delay in a core negative feedback loop[7,13,19]. Theoretical examinations of these oscillators predict that changes to the transcriptional delay (and hence the signaling time) can alter their period[20-22]. Similarly, transcriptional positive feedback loops can lead to bistable gene expression patterns[4,23] and theory predicts that the stability of the competing states increases as transcriptional signaling times increase[24].

Transcriptional signaling times can also impact the function of circuits that do not have feedback loops. Feedforward loops consist of parallel signaling pathways of different lengths that converge on the same transcriptional target[25]. The difference in the signaling times of those two pathways can alter the behavior of the target. For instance, incoherent feedforward loops can act as temporal pulse generators, and the pulse width will depend on the difference in timing of the two paths. Similarly, the behavior of any type of transcriptional circuit that relies on the temporal coincidence of multiple signals (such as transcriptional logic gates[26,27]) will depend on the respective signaling times of the upstream transcription factors.

Transcriptional signaling involves many sub-steps, including transcription and translation of the TF gene and its message, protein folding and oligomerization, the

accumulation of sufficient protein concentration, and promoter searching and binding. Each of these steps takes time and is affected by dilution, degradation, and cellular noise[28,29]. The timescales of some of the steps required for transcriptional signaling have been measured. For example, the transcription rate of RNA polymerase[30], translation rate of ribosome[31], and the promoter-searching rate of TF along DNA[32] have been characterized. However, the overall timescale and variability of transcriptional signaling times have not yet been determined. Such knowledge would aid in the computational design of synthetic gene circuits by providing needed dynamical information important for mathematical models[18].

Here we measured the transcriptional signaling time of two transcription factors in synthetic gene circuits. Specifically, we used microfluidics and time-lapse fluorescence microscopy[33,34] to quantify the mean and variability of signaling times of two key transcription factors used by synthetic biologists in *E. coli*: LacI and AraC. We found that average signaling times for both AraC and LacI were on the order of a few minutes at high induction levels. When induction levels were low, the mean and standard deviation of the signaling times greatly increased. Additionally, by examining the dynamics of a type-1 incoherent feedforward loop[25] and its sub circuits, we found that the signaling time of composite systems could be predicted from the signaling times of the constituent parts. This type of prediction is important for the forward engineering of ever more complex synthetic gene circuits[35,36].

Results

We first set out to experimentally characterize the transcriptional signaling time of a transcriptional activator regulating a downstream target. To do this, we built a plasmid-borne "activation circuit," depicted in Figure 1A. Specifically, we placed the gene encoding the transcriptional activator AraC (with the LAA version of the *ssrA* proteolysis tag[37]) under control of a LacI regulated promoter, $P_{A1lacO1,}$ and the gene encoding superfolding yellow fluorescent protein (sfYFP)[38] under control of the P_{BAD} promoter. In the absence of isopropyl β-D-1-thiogalactopyranoside (IPTG), the $P_{A1lacO1}$ promoter is repressed by a genomically integrated, constitutively expressed copy of *lacI*. In the presence of arabinose, the P_{BAD} promoter is up-regulated by AraC. Therefore, we could initiate the transcriptional signal by simultaneously introducing IPTG and arabinose into a microfluidic device in which cells containing the circuit were growing and measure the resulting output of the target promoter via YFP fluorescence (see Supporting Information). By varying the amount of IPTG (with arabinose held constant at 2% (w/v)), we could tune the mean level of downstream protein production (Fig. 1B). In addition, we also directly measured the "observation time" of sfYFP, *i.e*. the time it takes to first observe fluorescence without the need for the accumulation of AraC, by examining a simple "reporter only" circuit in which the P_{BAD} promoter drives *sfyfp* upon introduction of arabinose (Fig. S1C).

To measure the observation time of sfYFP, we transformed the reporter only circuit into the strain JS006A, which is a Δ*lacI*Δ*araC* cell strain with constitutively expressed *araC* knocked into the genome. We loaded cells into a microfluidic device capable of rapidly switching between two different sources of media. Cells were monitored via microfluidic aided fluorescence microscopy for 10 minutes in LB media before switching to media with 2% (w/v) arabinose. The fluorescence trajectories of single cells were recorded and the experiment was repeated to obtain at least 60 single-cell trajectories. To determine the time at which downstream fluorescence was first observed, we defined a threshold level of fluorescence 4 standard deviations above the fluorescence observed in non-inducing media (Fig. 2A). This threshold was determined to be the minimum threshold that accurately reflected the start of the increase in sfYFP fluorescence while not underestimating the observation time (see SI). The time at which the fluorescence of each cell reaches threshold was recorded (Fig. 2B), and the resulting histogram of maturation times is shown in Fig. 2C. From these measurements, we determined that the mean observation time of sfYFP to be 6.4±1.3 minutes.

To determine the signaling time of AraC activation, we transformed the activation circuit into the strain JS006LT, which is a Δ*lacI*Δ*araC* cell strain with constitutively expressed *lacI* and *tetR* knocked into the genome. We next measured the response of the circuit by inducing *araC* expression with three different IPTG concentrations: 0.05 mM, 0.2 mM, and 2

mM. These concentrations of IPTG were chosen to give low, medium, and high levels of induction, respectively, based on the measured activity of $P_{A1lacO1}$ as a function of IPTG (Fig. 1B). Specifically, cells were loaded into the microfluidic device as before, grown in LB media before switching to media containing varying amount of IPTG and 2% arabinose. Individual cells were monitored as described above, and their activation times were recorded (based on the threshold method described above).

For the activation circuit, when the inducer concentration is high (2mM IPTG), the estimated signaling time was near the observation time of sfYFP, namely 7.2±1.4 minutes (Fig. 2D-F). We hypothesized that reducing the amount of IPTG, and hence reducing the induction level of AraC, would increase the activation time as it would likely take longer to reach a sufficient concentration of AraC within the cell to activate the downstream promoter. Indeed, we found that reducing the IPTG concentration lead to an increase in signaling time and its variability: the signaling time increased to 13.9±4.1 minutes when the IPTG concentration was reduced to 0.2 mM, and to 27±13 minutes at 0.05 mM IPTG (Fig. 2G-L). Note that at 0.05 mM IPTG the steady-state expression level of sfYFP was very low, yet fluorescence increased discernably over background levels, and crossed our set threshold level (Fig. 2J).

We next tested whether signaling times under transcriptional repression exhibited similar behavior. To do so, we measured the signaling time of LacI. To build the “repression

circuit," we placed the gene encoding LacI (with the LAA tag) under control of the arabinose-inducible promoter P_{BAD} and the gene encoding *sfyfp* (with the LAA tag) under control of the hybrid $P_{lac/ara-1}$ promoter[39], which is activated by AraC and repressed by LacI (Fig. 3A). AraC was provided by a constitutively expressed genomic copy of *araC*. The circuit was designed so that *sfyfp* and *lacI* were activated simultaneously when arabinose was added. Once sufficient LacI accumulates, it should repress *sfyfp*. We therefore expected the circuit to act as a pulse generator with sfYFP fluorescence increasing, reaching a maximum, and then decaying once its production has been repressed. The strength and duration of the pulse are determined by the promoter activities of P_{BAD} and $P_{lac/ara-1}$, as determined by the arabinose concentration (Fig. 3B).

We transformed the plasmid containing the repression circuit into the strain JS006A, described above. We monitored cells as described above, but we instead triggered the circuit by introducing 4 different levels of arabinose: one with a high level of induction (2% w/v), one with a low level of induction (0.02%), and two different intermediate concentrations (0.1% and 0.05%). Note that we chose two different intermediate concentrations because the two promoters (P_{BAD} and $P_{lac/ara-1}$) respond differently to arabinose (Fig. 3B). Under each condition, we recorded at least 60 single-cell trajectories. We estimated the signaling time of LacI by measuring the time from the introduction of inducer to the time of peak sfYFP fluorescence. At high inducer concentration (2% arabinose), the signaling time was 8.6±1.1 minutes (Figure

4A-C). As in the activation circuit, the signaling time of the repression circuit and its variability increased with decreasing induction (Fig. 4D-L).

While the time to reach peak fluorescence is one measure of the signaling time in the repression circuit, it does not exactly reflect the minimum time for LacI to influence the downstream promoter. This is because the rate of sfYFP accumulation depends on both the production rate and the degradation rate of sfYFP. Peak fluorescence occurs when the degradation rate equals the production rate. However, LacI can still affect sfYFP production before this peak by reducing the production rate to a level that is still above the degradation rate. To estimate the minimal time of signaling, we examined the time derivative of the fluorescence trajectories. The peak of the derivatives indicates the inflection point in the recorded fluorescence, and thus provides an estimate of the minimal time it takes LacI to affect the target promoter. As shown in Fig. 5, this estimate of the signaling time is tightly distributed around 4 minutes at 2% arabinose (4.1±0.4 minutes). Note that this estimate provides an upper bound for the transcriptional delay time[16] (the time it takes to make one fully functional protein). Our measured time is consistent with theoretical estimates of the transcriptional delay time for LacI[20]. The minimal delay is thought to be necessary for the robustness of some synthetic genetic oscillators[7].

Finally, we wanted to determine if the dynamics of a circuit consisting of multiple transcriptional signaling pathways could be predicted from the signaling times of the

individual pathways in isolation. To do so we constructed a type-1 incoherent feedforward loop[25] shown in Fig. 6A. Specifically, the gene encoding *araC* was placed under the control of the TetR responsive promoter P_{tet} (which can be induced by anhydrotetracycline, ATc); *lacI* (with the LAA tag) was placed under the control of P_{BAD}, and *sfyfp* (with the LAA tag) under the control of $P_{lac/ara-1}$. Therefore, upon introduction of inducers, AraC was produced first and up-regulated both *lacI* and *sfyfp*. Once a sufficient concentration of LacI was reached within the cell, *sfyfp* was turned off and sfYFP fluorescence began to decrease. Note that this feedforward loop consists of two different paths. The short path is similar to the activation circuit described above. Here, we are more interested in the longer path (containing an initial activation step and a subsequent repression step) and the time it takes for it to turn off production of sfYFP.

We hypothesized that the dynamics of the feedforward loop could be predicted from the signaling times of its constituent pathways. Specifically, we wanted to estimate the time it takes to turn off production of sfYFP from the signaling times of the two steps in the longer pathway. To do so we first characterized the dynamics of three sub-circuits of the feedforward loop (Fig. 6): 1) the initial activation step (with signaling time T_1); 2) the subsequent repression step (with signaling time T_2); and 3) a small circuit that simply induces sfYFP:LAA (with signaling T_y). We expected that the signaling time of the combined long pathway, combining the activation and repressive steps, to equal $T_{TOT}=T_1+T_2-T_y$. We note that T_y must

be subtracted because the accumulation time of the intermediate protein (LacI in the feedforward loop) is measured twice – both as part of T_1 and T_2. Because we cannot measure the accumulation of LacI directly in single cells, we used sfYFP as a surrogate.

To test our hypothesis, we individually transformed each of the sub-circuits and the full feedforward loop into the strain JS006T, which has constitutively expressed *tetR* knocked into the genome, and assayed the cells as described above, using 1μg/ml ATc and 2% arabinose to induce the circuits. T_2 and T_y were measured above by the repression circuit and the P_{BAD} reporter only circuit (Fig. 6). The feedforward loop acts as a pulse generator with the peak fluorescence at 12.8±1.3 minutes after induction.

To obtain a prediction of the time at which the peak fluorescence occurs, we assumed that T_1, T_2, and T_Y are thee independent random variables. To estimate the distribution of T_{TOT} we sampled a value of $T_{1,i}$, $T_{2,i}$, and $T_{Y,i}$ from the measured values and set $T_{TOT,i}=T_{1,i}+T_{2,i}-T_{Y,i}$. We repeated this process 100,000 times. As shown in Fig. 6C, the distribution of the time at which fluorescence begins to turn off is close to the predicted distribution of T_{TOT} (13.3±2.4 minutes). However, we note that the standard deviation of the estimated time is slightly larger than the experimentally measured standard deviation. This might be due to correlations among the individual signaling times within the feedforward loop[40]. For instance, if AraC and LacI compete for the same resource, such as free ribosome or RNA polymerase, then T_1 and T_2 can be negatively correlated. We found that if these

correlations are taken into account, the standard deviation of the estimated distribution of T_{TOT} is closer to the experimentally measured standard deviation (see SI).

Discussion

We measured the transcriptional signaling time for the activator AraC and the repressor LacI, two transcription factors commonly used in the construction of synthetic genetic circuits. When they are highly expressed, transcriptional signaling times can be as short as a few minutes. Not unexpectedly, reducing the expression rate of a TF leads to an increased signaling time. When a TF is weakly expressed, the TF concentration inside the cell stays below the threshold needed to regulate its target promoter effectively[41], and the signaling time becomes highly variable for single cells – a fact that can be explained with simple gene regulation models[42]. Note the measured signaling times here include the ligand uptake time, which may become significant at low concentration[43]. However, it is difficult to measure the ligand uptake time using a fluorescent protein reporter, since ligand concentration affects both the uptake time and the level of target gene expression. Nevertheless, the measurements here provide the upper limit for the delay of transcriptional regulation, and the possible range of this delay when the production rate of a TF is varied.

We also examined a two-step process to determine if the timing and variability of each step can be used to predict the timing and variability of the entire pathway. Interestingly, while the mean signaling time of the entire cascade can be predicted from its constituent parts, its overall variability is smaller than that predicted from the simple convolution of the constituent distributions. It is possible that this could be explained by correlations in the signaling times of the two constituent pathways. If a cell can make the first transcription factor within the pathway quickly, it might make the second transcription factor slowly. In that case, the timing of the first step will be negatively correlated with the second step, making the variability of the two-step pathway smaller than the convolution of the individual signaling time distributions. Although the predicted distribution of the signaling time is slightly wider than the measured one, the similarity between the mean signaling time showed that our measurements are precise enough to predict a more complex circuit's signaling time.

Since the transcriptional signaling times are determined by the TF production rate, it is possible that one could fine-tune synthetic gene circuits with regulatory mechanisms such as CRISPR interference[44] or antisense RNA[45]. For example, it might be possible to extend the period of a synthetic gene oscillator by introducing additional delay into the negative-feedback loop[20]. It would also be interesting to see whether the findings here hold for other TFs. If other TFs behave similarly, a library of functional modules that have predictable dynamics can be built using different TFs.

As noted above, delays in transcriptional signaling can strongly affect the dynamics of circuits in which feedback loops, especially oscillators and bistable switches[20,24]. However, signaling times are difficult to measure in these circuits. Our approach allowed us to measure the signaling times of the constituent components of these circuits, and provided an upper bound of transcription delay consistent with theoretical estimates[20,46]. Moreover, we have also shown that the delay in multi-step pathways can be estimated from the constituent steps. Therefore, while the functions of the circuits we examined here are not drastically affected by signaling times, they do allow us to measure them. And while one cannot *a priori* assume that the signaling times will be similar in other types of circuits, these measurements at least provide an understanding of the relevant time scales. As a further caveat, it should be noted that these signaling times were measured on plasmid-borne synthetic circuits in bacteria. Circuits built in the chromosome of bacteria could behave differently due to the lower copy numbers of the constituent genes[47], and circuits built in eukaryotes will have additional complicating factors such as nuclear localization of transcription factors and chromatin remodeling[48].

Acknowledgments

This work was funded by the National Institutes of Health, through the joint NSF/NIGMS mathematical biology program grant R01GM104974 (MRB and KJ), the National Science Foundation grant DMS-1122094 (KJ), the Welch Foundation grant C-1729 (MRB), and the Taiwan Studying Abroad Scholarship (YYC). MRB, YYC, and KJ conceived and designed the study. YYC performed the experiments, analyzed the data and ran the computational simulations. AJH designed and fabricated the microfluidic device. MRB supervised the project. All authors wrote the manuscript.

Supporting Information

Materials and Methods, supplementary figures.

References

1 Mutalik, V. K. *et al.* (2013) Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat Meth* **10**, 354-360.

2 Prindle, A. *et al.* (2014) Rapid and tunable post-translational coupling of genetic circuits. *Nature* **508**, 387-391.

3 Qi, H., Blanchard, A. & Lu, T. (2013) Engineered genetic information processing circuits. *Wiley Interdiscip Rev Syst Biol Med* **5**, 273-287.

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

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

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

7 Stricker, J. *et al.* (2008) A fast, robust and tunable synthetic gene oscillator. *Nature* **456**, 516-519.

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

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

10 Nevozhay, D., Adams, R. M., Murphy, K. F., Josic, K. & Balazsi, G. (2009) Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression. *Proc Natl Acad Sci USA* **106**, 5123-5128.

11 Brenner, K., Karig, D. K., Arnold, F. H. & Weiss, R. (2007) Engineered bidirectional communication mediates a consensus in a microbial biofilm consortium. *Proc Natl Acad Sci USA* **104**, 17300-17304.

12 Balagadde, F. K. *et al.* (2008) A synthetic Escherichia coli predator-prey ecosystem. *Mol Syst Biol* **4**, 187.

13 Danino, T., Mondragon-Palomino, O., Tsimring, L. & Hasty, J. (2010) A synchronized quorum of genetic clocks. *Nature* **463**, 326-330.

14 Chen, Y., Kim, J. K., Hirning, A. J., Josic, K. & Bennett, M. R. (2015) Emergent genetic oscillations in synthetic microbial consortia. *Science* **349**, 986-989.

15 Nielsen, A. A. K. *et al.* (2016) Genetic circuit design automation. *Science* **352**, 53.

16 Gupta, C. *et al.* (2014) Modeling delay in genetic networks: From delay birth-death processes to delay stochastic differential equations. *J Chem Phys* **140**, 204108.

17 Note that here we use the terms “transcriptional signaling” and “signaling times” in relation to the process of a TF being first made and then signaling its downstream target promoter. This process should not be confused with signaling due to post-translational modifications.

18 Bennett, M. R., Volfson, D., Tsimring, L. & Hasty, J. (2007) Transient dynamics of genetic regulatory networks. *Biophys J* **92**, 3501-3512.

19 Hussain, F. *et al.* (2014) Engineered temperature compensation in a synthetic genetic clock. *Proc Natl Acad Sci USA* **111**, 972-977.

20 Mather, W., Bennett, M. R., Hasty, J. & Tsimring, L. S. (2009) Delay-Induced Degrade-and-Fire Oscillations in Small Genetic Circuits. *Phys Rev Lett* **102**, 068105.

21 Josic, K., Lopez, J. M., Ott, W., Shiau, L. & Bennett, M. R. (2011) Stochastic Delay Accelerates Signaling in Gene Networks. *PLoS Comp Biol* **7**, E1002264.

22 O'Brien, E. L., Van Itallie, E. & Bennett, M. R. (2012) Modeling Synthetic Gene Oscillators. *Math Biosci* **236**, 1-15.

23 Egbert, R. G. & Klavins, E. (2012) Fine-tuning gene networks using simple sequence repeats. *Proc Natl Acad Sci USA* **109**, 16817-16822.

24 Gupta, C., Lopez, J. M., Ott, W., Josic, K. & Bennett, M. R. (2013) Transcriptional Delay Stabilizes Bistable Gene Networks. *Phys Rev Lett* **111**, 058104.

25 Mangan, S. & Alon, U. (2003) Structure and function of the feed-forward loop network motif. *Proc Natl Acad Sci USA* **100**, 11980-11985.

26 Stanton, B. C. *et al.* (2014) Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat Chem Biol* **10**, 99-105.

27 Shis, D. L. & Bennett, M. R. (2013) Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants. *Proc Natl Acad Sci USA* **110**, 5028-5033.

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

29 Gronlund, A., Lotstedt, P. & Elf, J. (2010) Costs and constraints from time-delayed feedback in small gene regulatory motifs. *Proc Natl Acad Sci USA* **107**, 8171-8176.

30 Vogel, U. & Jensen, K. F. (1994) The RNA Chain Elongation Rate in Escherichia coli Depends on the Growth-Rate. *J Bacteriol* **176**, 2807-2813.

31 Proshkin, S., Rahmouni, A. R., Mironov, A. & Nudler, E. (2010) Cooperation Between Translating Ribosomes and RNA Polymerase in Transcription Elongation. *Science* **328**, 504-508.

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

33 Bennett, M. R. & Hasty, J. (2009) Microfluidic devices for measuring gene network dynamics in single cells. *Nat Rev Genet* **10**, 628-638.

34 Ferry, M. S., Razinkov, I. A. & Hasty, J. (2011) Microfluidics for Synthetic Biology: From Design to Execution. *Method Enzymol* **497**, 295-372.

35 Murphy, K. F., Balazsi, G. & Collins, J. J. (2007) Combinatorial promoter design for engineering noisy gene expression. *Proc Natl Acad Sci USA* **104**, 12726-12731.

36 Guido, N. J. *et al.* (2006) A bottom-up approach to gene regulation. *Nature* **439**, 856-860.

37 Andersen, J. B. *et al.* (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl Environ Microbiol* **64**, 2240-2246.

38 Kremers, G. J., Goedhart, J., van Munster, E. B. & Gadella, T. W. J. (2006) Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Forster radius. *Biochemistry* **45**, 6570-6580.

39 Lutz, R. & Bujard, H. (1997) Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I-1-I-2 regulatory elements. *Nucl Acids Res* **25**, 1203-1210.

40 Pedraza, J. M. & Paulsson, J. Random timing in signaling cascades. (2007) *Mol. Syst. Biol.* **3**, 81.

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

42 Ghusinga, K. R., Dennehy, J. J. & Singh, A. (2017) First-passage time approach to controlling noise in the timing of intracellular events. *Proc Natl Acad Sci USA* **114**, 693-698.

43 Fritz, G. *et al.* (2014) Single Cell Kinetics of Phenotypic Switching in the Arabinose Utilization System of E. coli. *Plos One* **9**, e89532.

44 Qi, L. S. *et al.* (2013) Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**, 1173-1183.

45 Hoynes-O'Connor, A. & Moon, T. S. (2016) Development of Design Rules for Reliable Antisense RNA Behavior in E. coli. *ACS Synth Biol* **5**, 1441-1454.

46 Veliz-Cuba, A. *et al.* (2015) Sources of Variability in a Synthetic Gene Oscillator. *PLoS Comput Biol* **11**, e1004674.

47 Mileyko, Y., Joh, R. I. & Weitz, J. S. (2008) Small-scale copy number variation and large-scale changes in gene expression. *Proc Natl Acad Sci USA* **105**, 16659-16664.

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

Figure Captions

Figure 1. Activation circuit and the characterization of $P_{A1lacO1}$ promoter. (A) The activation circuit was used to measure the signaling time of *araC*. Once IPTG and arabinose are added, *araC* will be induced to activate *yfp*. The YFP fluorescence signal was used to determine the signaling time. (B) Relative $P_{A1lacO1}$ promoter activity (measured as fluorescence divided by OD_{600}) as a function of IPTG concentration. From this curve, we selected 3 concentrations of IPTG with which to test the activation circuit: 0.05 mM, 0.2 mM, and 2 mM.

Figure 2. Single-cell trajectories of the P_{BAD} reporter only circuit and the activation circuit and the estimated signaling times. (A) Example single-cell trajectory of the reporter only circuit. Shown are the fluorescence as a function of time (black curve), and the threshold value used to determine the time at which fluorescence was first observed (red line). (B) All single cell fluorescence trajectories (black curves) of the reporter only circuit. Also shown is the example trajectory from panel A (red curve). (C) Histogram of the measured observation times of the reporter only circuit. (D) Example single-cell trajectory of the activation circuit induced with 2mM IPTG. Shown are the fluorescence as a function of time (black curve), and the threshold value used to determine the time at which fluorescence was first observed (red line). (E) All single cell fluorescence trajectories (black curves) of the activation circuit induced with 2mM IPTG. Also shown is the example trajectory from panel D (red curve). (F) Histogram of the measured signaling times of the activation circuit induced with 2mM IPTG. (G)-(I) Same as panels D-F, but induced with 0.2 mM IPTG. (J)-(L) Same as panels D-F, but induced with 0.05 mM IPTG.

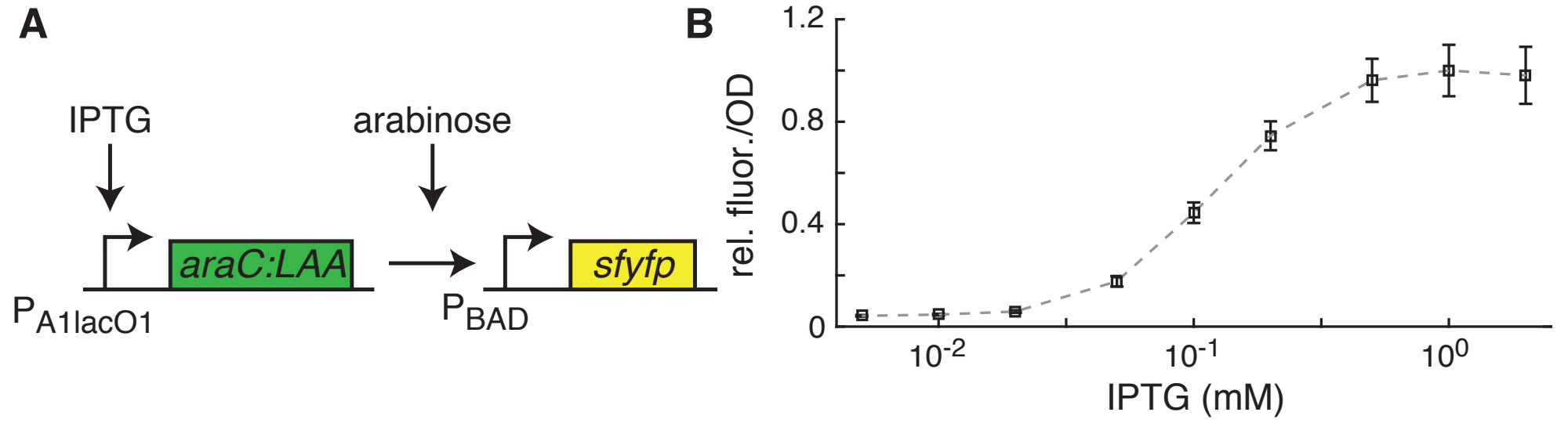
Figure 3. Repression circuit and the characterization of P_{BAD} and $P_{lac/ara-1}$ promoters. (A) The repression circuit was used to measure the signaling time for *lacI* to repress the $P_{lac/ara-1}$ promoter. Once arabinose is added, constitutively expressed AraC will activate both *lacI* and *yfp*. Once enough functional LacI has accumulated, *yfp* expression will be repressed. (B)

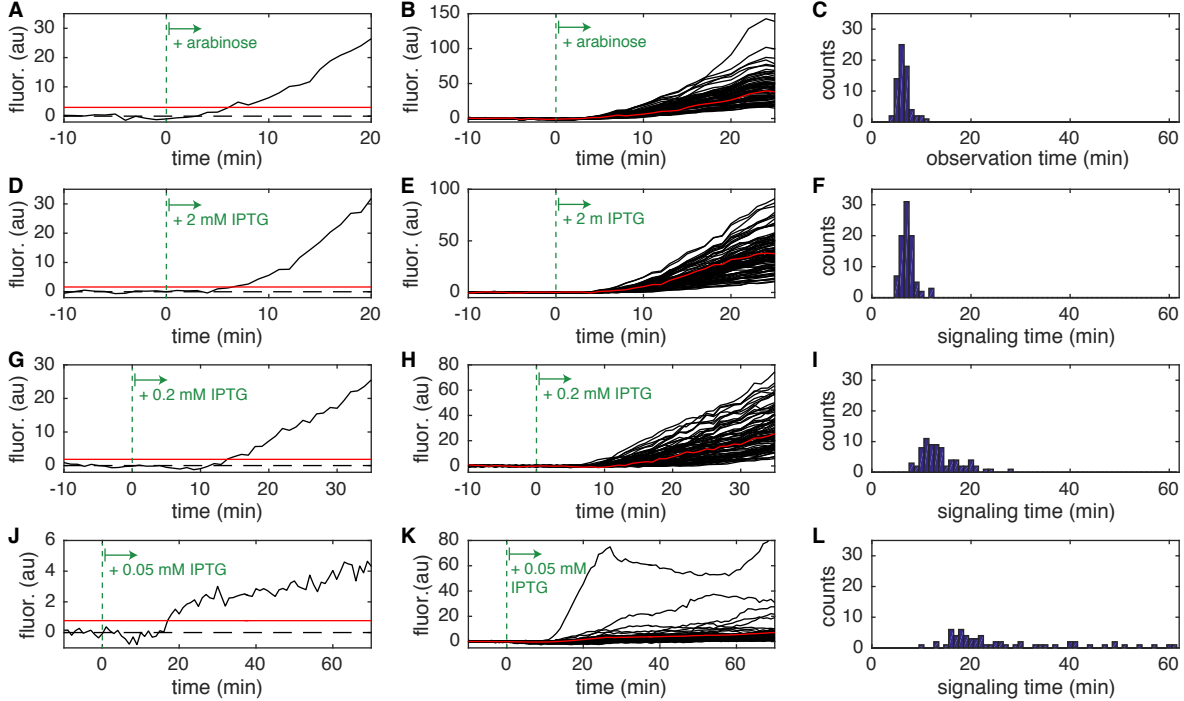
Relative activity of the P_{BAD} (black squares) and $P_{lac/ara-1}$ (red circles) promoters (measured as fluorescence divided by OD_{600}) as a function of arabinose concentration. From these curves, we chose to induce the repression circuit 0.02%, 0.05%, 0.1%, and 2% arabinose.

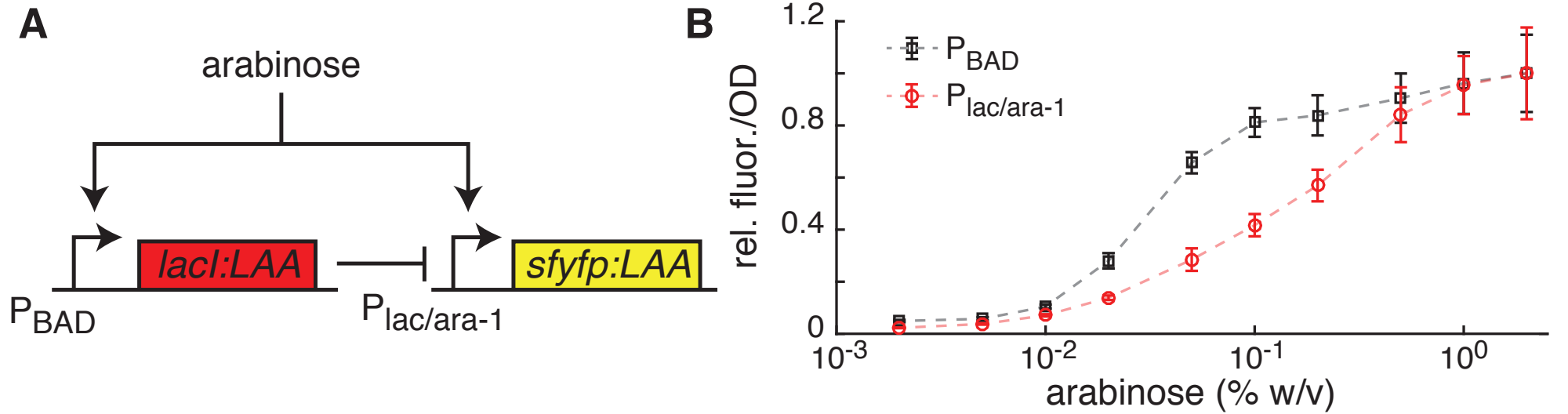
Figure 4. Single-cell trajectories of the repression circuit and the estimated signaling times. (A) Example single-cell trajectory of the repression circuit induced with 2% arabinose. Shown is the fluorescence as a function of time. (B) All single cell fluorescence trajectories (black curves) of the repression only circuit induced with 2% arabinose. Also shown is the example trajectory from panel A (red curve). (C) Histogram of the measured signaling times of the repression circuit induced with 2% arabinose. (D)-(F) Same as panels A-C, except induced with 0.1% arabinose. (G)-(I) Same as panels A-C, except induced with 0.05% arabinose. (J)-(L) Same as panels A-C, except induced with 0.02% arabinose.

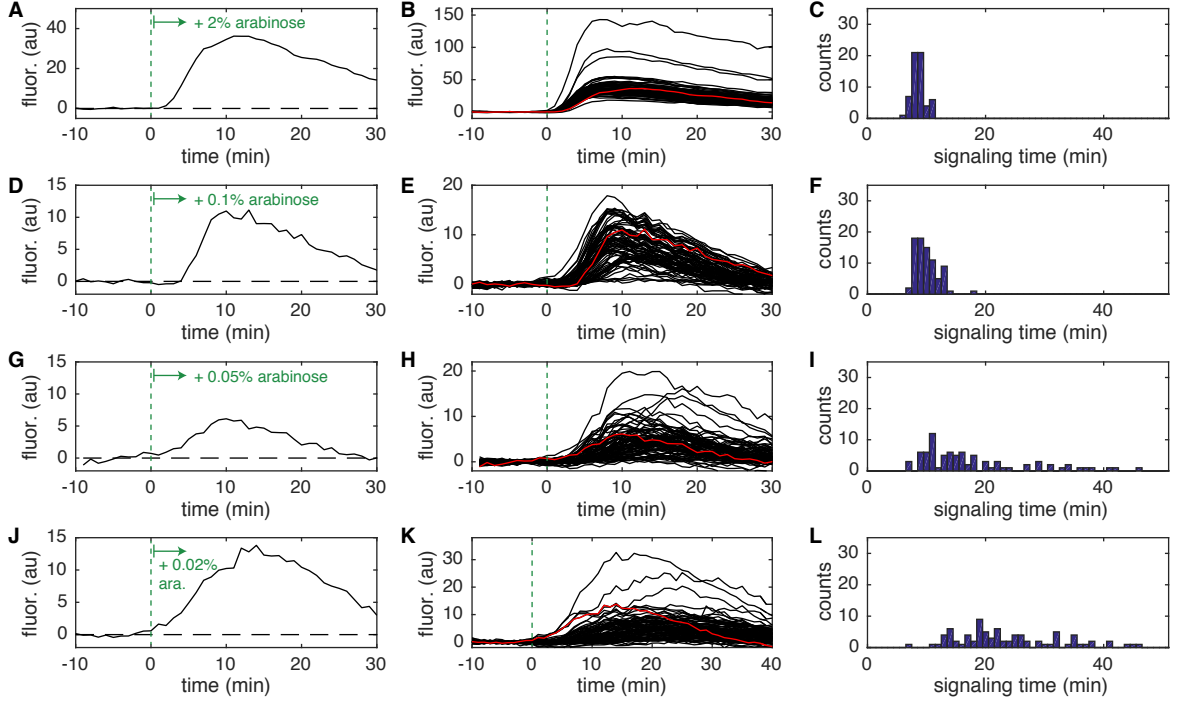
Figure 5. Estimation of the minimal LacI signaling time. (A) Time derivatives of the single-cell trajectories of the repression circuit induced with 2% ARA (Fig. 4B). (B) Histogram of the minimal LacI signaling times, estimated by the peak of the production rate trajectories shown in panel A. The estimated minimal LacI signaling time is 4.1±0.4 minutes.

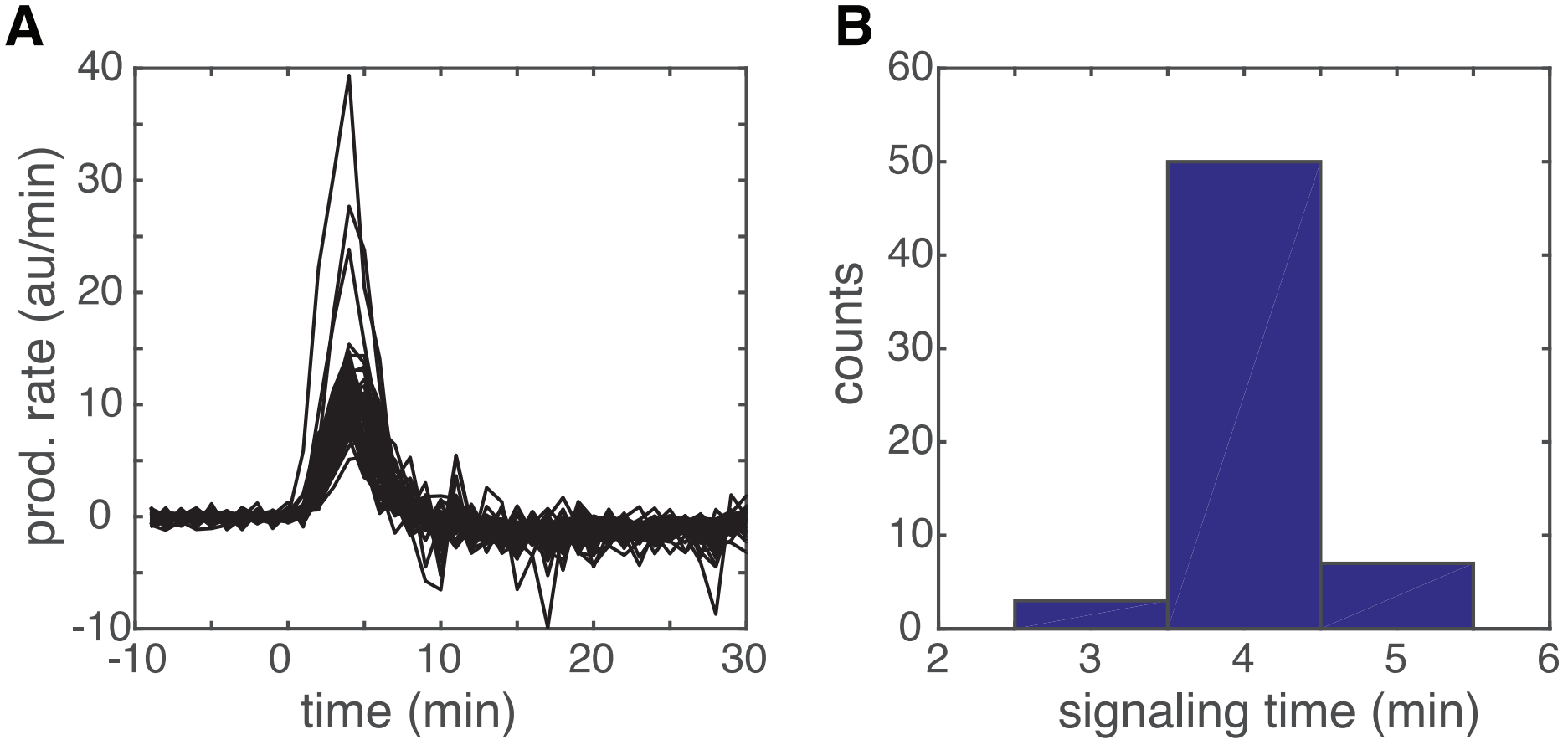
Figure 6. Prediction of the signaling time of the feedforward loop. (A) Circuit diagram of the incoherent feedforward loop. (B) Simplified circuit schematic highlighting the signaling time of interest (in gray). This signaling time, T_{TOT}, is the summation of the time for AraC to activate l*acI* and the time for LacI to repress *yfp*. (C) Histogram of the experimentally measured T_{TOT} of the two-step cascade (blue bars). Red curve is the predicted signaling time, which was convolved from the experimental measurements of the three sub-circuits (below). (D) Circuit schematic of the modified activation circuit. (E) The modified activation circuit was used to measure the signaling time T_1, which is the time for AraC to activate LacI, plus the YFP observation time. (F) Histogram of the measured T_1 of the modified activation circuit. (G) Circuit schematic of the repression circuit. (H) The repression circuit was used to measure the signaling time T_2, which is the time for LacI to repress YFP. (I) Histogram of the measured T_2 of the repression circuit (from Fig. 4C). (J) Circuit schematic of the reporter only circuit. (K) The P_{BAD} reporter only circuit was used to measure YFP observation time, T_y. (L) Histogram of the measured T_y of the reporter only circuit (from Fig. 2C).

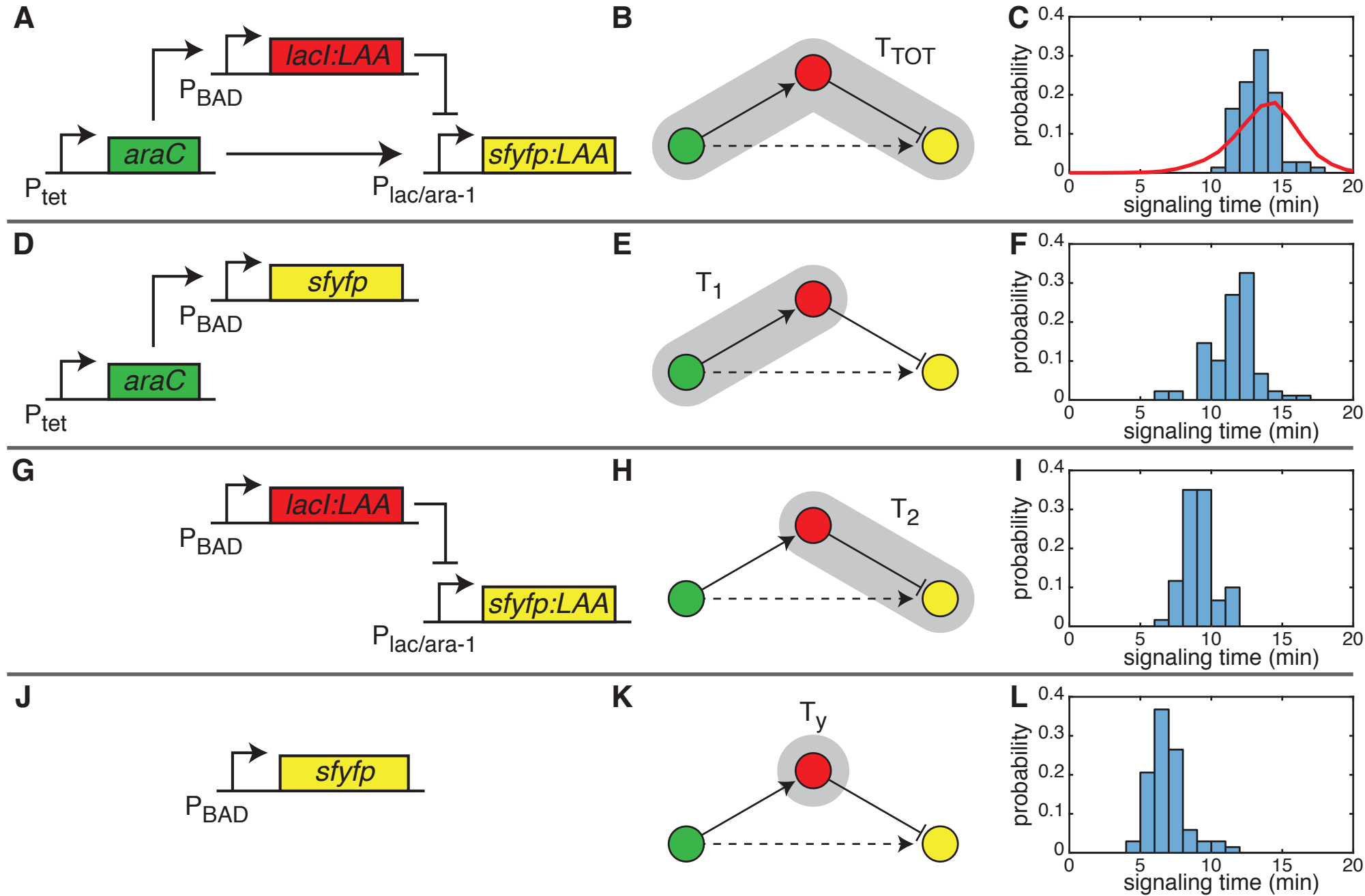

A
IPTG
arabinose
araC:LAA
sfyfp
P A1lacO1
P BAD
B
rel. fluor./OD
1.2
0.8
0.4
0
10-2
10-1
100
IPTG (mM)


A
fluor. (au)
+ arabinose
time (min)
B
fluor. (au)
+ arabinose
time (min)
C
counts
observation time (min)
D
fluor. (au)
+ 2 mM IPTG
time (min)
E
fluor. (au)
+ 2 m IPTG
time (min)
F
counts
signaling time (min)
G
fluor. (au)
+ 0.2 mM IPTG
time (min)
H
fluor. (au)
+ 0.2 mM IPTG
time (min)
I
counts
signaling time (min)
J
fluor. (au)
+ 0.05 mM IPTG
time (min)
K
fluor.(au)
+ 0.05 mM IPTG
time (min)
L
counts
signaling time (min)


A
arabinose
lacI:LAA
sfyfp:LAA
P BAD
P lac/ara-1
B
rel. fluor./OD
1.2
0.8
0.4
0
P BAD
P lac/ara-1
10-3
10-2
10-1
100
arabinose (% w/v)


A
+ 2% arabinose
fluor. (au)
time (min)
B
fluor. (au)
time (min)
C
counts
signaling time (min)
D
+ 0.1% arabinose
fluor. (au)
time (min)
E
fluor. (au)
time (min)
F
counts
signaling time (min)
G
+ 0.05% arabinose
fluor. (au)
time (min)
H
fluor. (au)
time (min)
I
counts
signaling time (min)
J
+ 0.02% ara.
fluor. (au)
time (min)
K
fluor. (au)
time (min)
L
counts
signaling time (min)


A
40
30
20
10
0
-10
prod. rate (au/min)
-10
0
10
20
30
time (min)
B
60
50
40
30
20
10
0
counts
2
3
4
5
6
signaling time (min)


A
lacI:LAA
P BAD
araC
sfyfp:LAA
P tet
P lac/ara-1
B
T TOT
C
probability
signaling time (min)
D
sfyfp
P BAD
araC
P tet
E
T 1
F
probability
signaling time (min)
G
lacI:LAA
P BAD
sfyfp:LAA
P lac/ara-1
H
T 2
I
probability
signaling time (min)
J
sfyfp
P BAD
K
T y
L
probability
signaling time (min)

Supplementary Information:
The timing of transcriptional regulation in synthetic gene circuits

Yu-Yu Cheng[1], Andrew J. Hirning[1], Krešimir Josić[1,2,3], and Matthew R. Bennett[1,4,†]

[1]Department of Biosciences, Rice University, Houston, TX 77005, USA.
[2]Department of Mathematics, University of Houston, Houston, TX 77204, USA.
[3]Department of Biosciences, University of Houston, Houston, TX 77204, USA.
[4]Department of Bioengineering, Rice University, Houston, TX 77005, USA.
[†]Corresponding author. E-mail: matthew.bennett@rice.edu

Materials and Methods

Plasmids

All circuits were constructed on plasmids (Fig. S1). *araC* and *lacI* are tagged with the LAA version of the *ssrA* degradation tag[1] in all circuit designs, except for *araC* in the P_{Tet}-*araC* circuit (Fig. S1D and S1F). *sfyfp* is also tagged with the LAA version of *ssrA* in the repression circuit, the cascade, the P_{LlacO1}-*sfyfp* reporter-only circuit, and the $P_{lac/ara-1}$-*sfyfp* reporter-only circuit (Fig. S1B, S1D, S1E and S1G). We used a bicistronic ribosome binding site[2] (BCD) for all genes, except for *araC* in the activation circuit (Fig. S1A), for which we used the B0034 ribosome binding site (http://partsregistry.org/Part:BBa_B0034). All circuits were constructed on plasmids with the p15A origin, except for P_{Tet}-*araC*, which is on a plasmid with the pSC101 origin (Fig. S1D and S1F).

Strain construction

The three *E. coli* strains used in this study (JS006LT, JS006A, and JS006T) were derived from the JS006 strain (BW25113 *ΔlacIΔaraC*)[3]. The activation circuit (Fig. S1A) and P_{LlacO1}-*sfyfp* reporter-only circuit (Fig. S1E) were transformed into JS006LT, which contains constitutively expressed *lacI* and *tetR*. The repression circuit (Fig. S1B), P_{BAD}-*sfyfp* reporter-only circuit (Fig. S1C), and $P_{lac/ara-1}$-*sfyfp* reporter-only circuit (Fig. S1G) were transformed into JS006A, which contains constitutively expressed *araC*. The two-step genetic cascade (Fig. S1D) and the modified activation circuit (Fig. S1F) were transformed into JS006T, which contains constitutively expressed *tetR*. The genes were knocked into the genome via the lambda integrase method[4].

Microfluidic device and microscopy setup

We used microfluidic devices to observe single cells over time with fluorescence

microscopy. Master molds and microfluidic chips were created with standard photolithography and soft lithography techniques[5]. The 'dial-a-wave' chip was designed to fine-tune the culture medium constituents for the cells (Fig. S2A). This device was created by integrating the dial-a-wave junction from Bennett *et al.* 2008[6] with the microfluidic chip used in Hussain *et al.* 2014[7]. The dye sulforhodamine 101 (Sigma-Aldrich) was used (at 1 μg/ml) to track the concentration of the inducers. The typical dye trajectory of the step function (Fig. S2B) shows that it needs 5 minutes for the inducer to reach maximal concentration. We defined the first time point at which inducer started to increase as time 0. We found that for *yfp* expressed from strong promoters, such as $P_{lac/ara}$ promoter, the *yfp* signal can be detected within 1 to 2 minutes (Fig. S2C).

For each microscopy experiment, cells were first loaded into the 'dial-a-wave' chip. When about 30 cells were trapped in the cell trapping region, phase contrast (100x magnification with additional 1.5x magnification) and fluorescence images (mCherry and YFP channels) were taken every 1 minute. The power of the excitation light was 20% power of a 200W light source (Lumen 200 Pro). The exposure time was 300ms and 100ms for sfYFP and mCherry, respectively. Cells remained healthy after 8 hours of image acquisition. The inducers were loaded into the chip after 10 minutes. We used the 10-minute background YFP signal to define the threshold for the activation time. Each experiment was performed at least twice to get a minimum of 60 single-cell trajectories for the statistics.

Image analysis and data acquisition

Images were analyzed using a semi-automatic tracking algorithm[5] developed using MATLAB (Mathworks, Inc.). Cells were manually segmented using phase-contrast images. The results of the segmentations were used by the algorithm to record fluorescence at the regions identified as individual cells. The output fluorescence is the summed fluorescence intensity divided by the number of pixels in the segmented cell region. When the cells divided, we randomly picked one of the daughter cells to follow. The script is available for download online (github.com/alanavc/rodtracker).

Promoter characterization

Overnight cultures of the reporter-only strains – P_{BAD}-*yfp* reporter-only circuit in JS006A, P_{lac}-*yfp* reporter-only circuit in JS006LT, and $P_{lac/ara}$-*yfp* reporter-only circuit in JS006A, were diluted 1:1000 into 96-well plates with varying levels of inducers. The cells were then grown at 37°C for 2 hours. Fluorescence and OD_{600} were measured with an Infinite® 200 PRO fluorescence plate reader (TECAN). Each data point represents the average of

three measurements (Fig. 1B and 2A). Data is normalized because the $P_{lac/ara-1}$ promoter is approximately 10 times stronger than the P_{BAD} promoter.

Threshold for the estimation of signaling time of the activation circuit

To estimate the time at which the YFP signal starts to increase for the activation circuit, we defined the threshold to be the mean plus N standard deviations of the 10-minute background fluorescence signal, where N is a positive integer. We found that when N<4, the estimated signaling time for some cells is shorter than the YFP maturation time. Figure S3 shows the case for N=3, which lead to an estimate of less than one minute for the expression of YFP in some cells. The YFP maturation time was estimated using the reporter-only circuit and found to be at least 4 minutes. Therefore, we chose N=4 to avoid underestimating the signaling time.

Estimation of the signaling time in the two-step genetic cascade

The measured times T_1, T_y, and T_2 were used to estimate the total signaling time, T_{TOT}, in the cascade (Fig. 3C). We first assumed that each of the measured times are independent random variables. To estimate the distribution of the total signaling time, T_{TOT} (Fig. 3D), we sampled three times independently from the collection of measured values, $T_{1,i}$, $T_{y,i}$, and $T_{2,i}$, from the measured values of T_1, T_y, and T_2. We estimated the total signaling time from these three sampled values as $T_{TOT,i} = T_{1,i} - T_{y,i} + T_{2,i}$, and repeated the process 100,000 times (with replacement). The mean of distribution is very close (within 1 minute) to that found experimentally. However, the estimated standard deviation is slightly larger than what was measured (2.4 minutes versus 1.3 minutes). This might due to correlations between the three measurements. For example, *araC* and *lacI* expression could be negatively correlated since they compete for the same resources[8]. This will lead to a negative correlation between T_1 and T_2.

To take into account the possible correlation between T_1 and T_2, we first assume that they are normally distributed for simplicity. We fit the measurements of T_1 and T_2 assuming normality to obtain the mean μ_1 and μ_2 and standard deviation σ_1 and σ_2. Then T_{1+2} can be assumed to follow a normal distribution with mean μ_1 and standard deviation $\sqrt{{\sigma_1}^2 + {\sigma_2}^2 - 2\rho * \sigma_1 * \sigma_2}$, where ρ is the correlation coefficient between T_1 and T_2. The same Monte Carlo method was used to calculate T_{TOT} by subtracting T_y from the newly estimated T_{1+2}. When $\rho < 0$, the standard deviation of estimated T_{TOT} is smaller, with the mean not changed. Figure S4 shows the case for $\rho = -0.5$ for illustration.

References

1 Andersen, J. B. *et al.* New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl Environ Microbiol* **64**, 2240-2246 (1998).

2 Mutalik, V. K. *et al.* Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat Meth* **10**, 354-360 (2013).

3 Stricker, J. *et al.* A fast, robust and tunable synthetic gene oscillator. *Nature* **456**, 516-519 (2008).

4 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci USA* **97**, 6640-6645 (2000).

5 Veliz-Cuba, A. *et al.* Sources of Variability in a Synthetic Gene Oscillator. *PLoS Comput Biol* **11**, e1004674 (2015).

6 Bennett, M. R. *et al.* Metabolic gene regulation in a dynamically changing environment. *Nature* **454**, 1119-1122 (2008).

7 Hussain, F. *et al.* Engineered temperature compensation in a synthetic genetic clock. *Proc Natl Acad Sci USA* **111**, 972-977 (2014).

8 Qian, Y., Huang, H.-H., Jiménez, J. I. & Del Vecchio, D. Resource Competition Shapes the Response of Genetic Circuits. *Acs Synth Biol*, Epub ahead of print, doi:10.1021/acssynbio.6b00361 (2017).

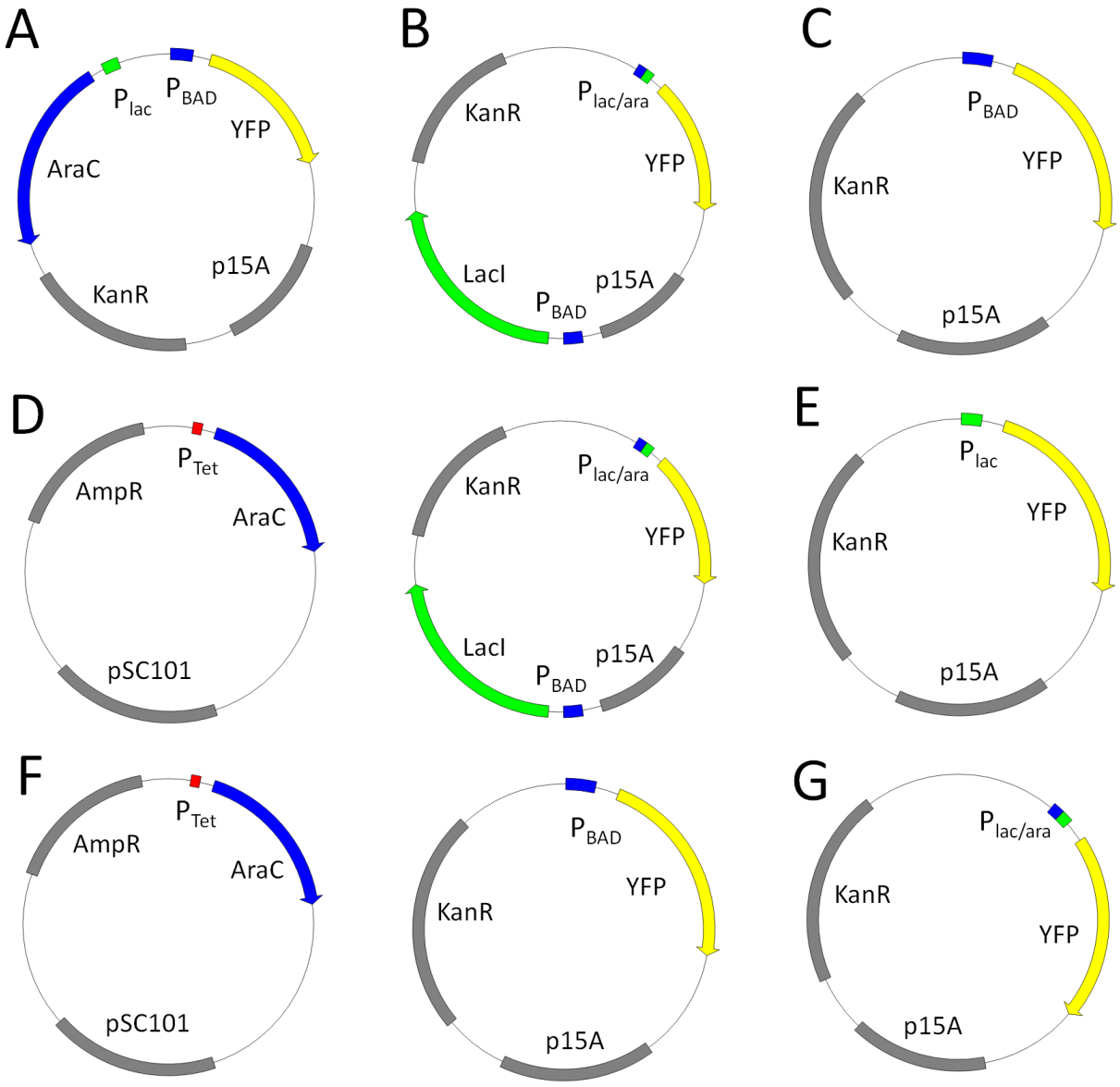


Figure. S1. Plasmids built for each genetic circuit. (A) Activation circuit. (B) Repression circuit. (C) P_{BAD} reporter only circuit. (D) Two-step genetic cascade. (E) P_{lac} reporter only circuit. (F) Modified activation circuit. (G) $P_{lac/ara}$ reporter only circuit. Plasmid maps were drawn with Savvy (http://bioinformatics.org/savvy/).

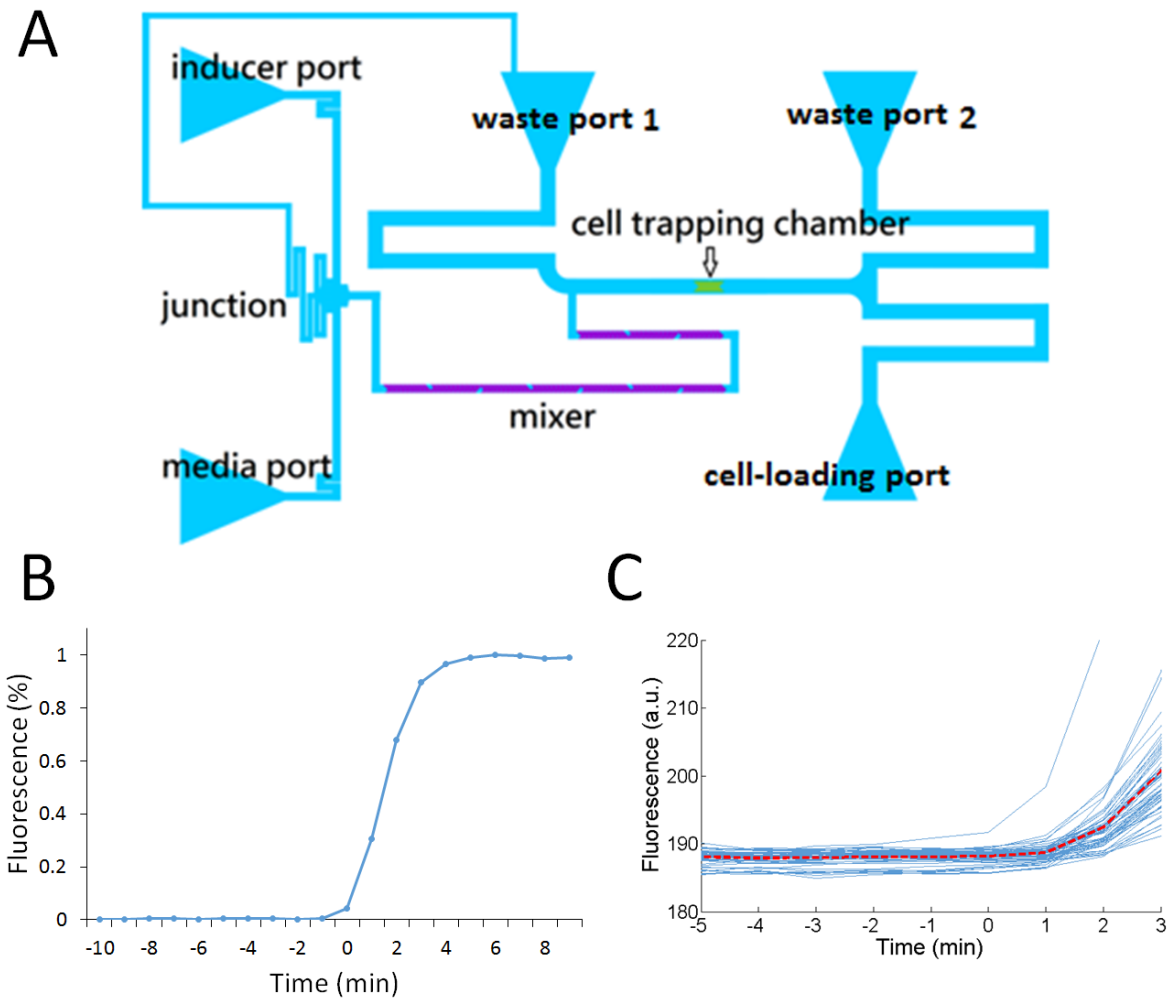


Figure. S2. The 'dial-a-wave' chip and its characterization. (A) The chip has 5 ports connected to external culture media. The media port is for LB media. The inducer port is for inducers and the dye Sulforhodamine 101. The two waste ports are for the outlet of the culture media. The junction and mixer are designed to fine-tune the inducer concentration with more precision. The cell-trapping chamber is 100 μm wide, 300 μm long, and 0.95 μm high. (B) The typical dye trajectory in the microfluidic experiments. It shows that inducers need 5 minutes to reach maximum concentration in the chamber after loading. Time 0 is defined at the point that the fluorescence starts to increase. (C) YFP trajectories for the repression circuit at 2% ARA (Figure 2B). YFP increase can be detected 1 to 2 minutes after inducer concentration starts to increase. The dashed red line is the average fluorescence.

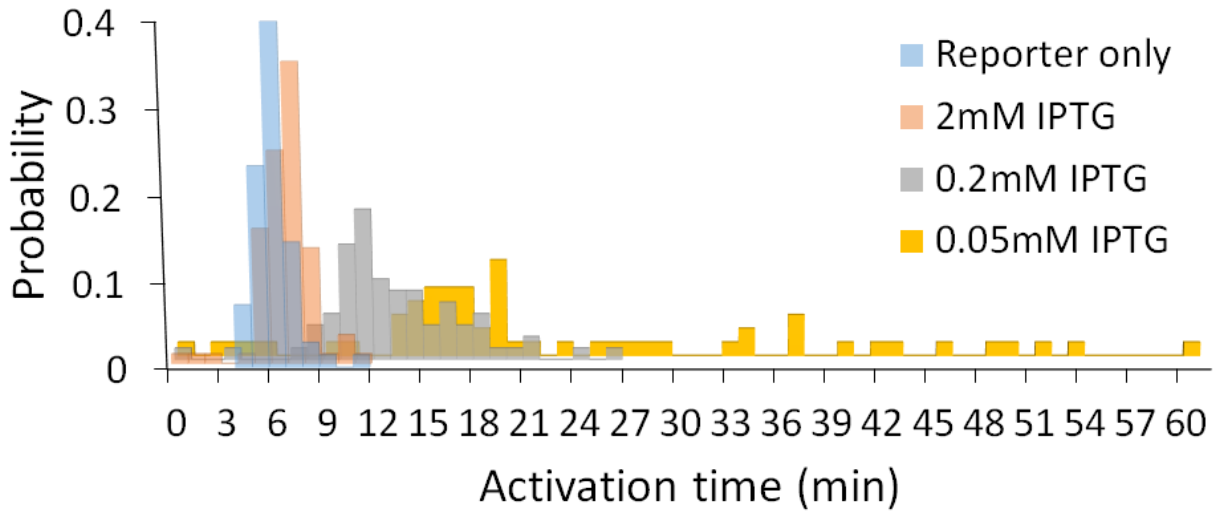


Figure S3. The estimated signaling time for the activation circuit when the threshold is defined as the mean plus three standard deviations of the background fluorescence signals. Some cells are estimated to express YFP in less than the estimated maturation time.

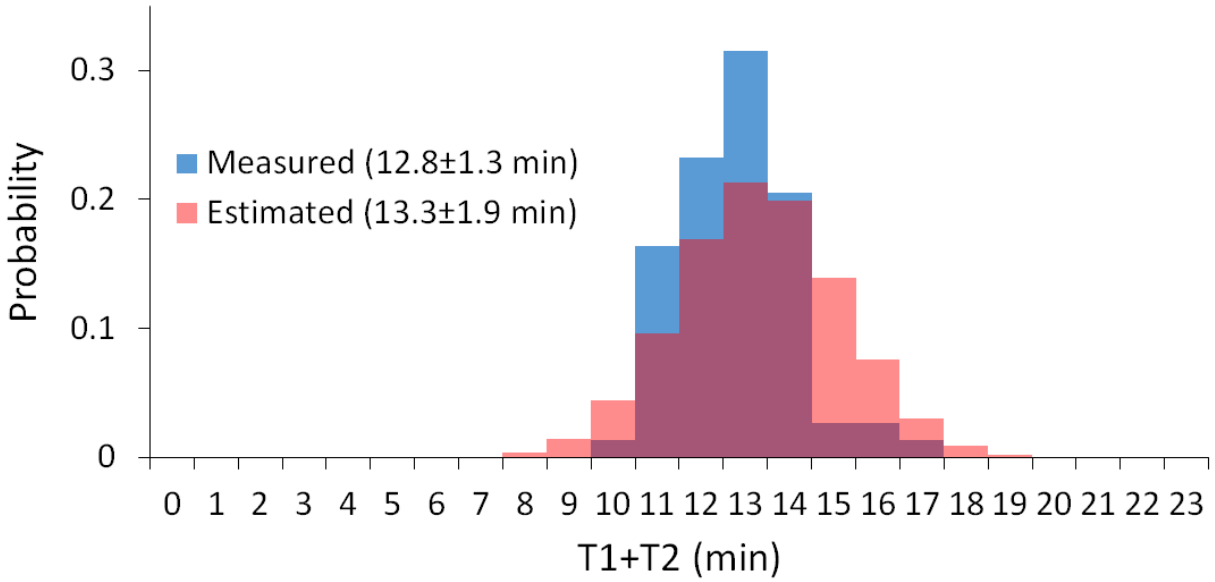


Figure S4. The comparison of the estimated signaling time with the measured signaling time of the cascade when T_1 and T_2 have correlation coefficient -0.5. The standard deviation of the estimated distribution is reduced from 2.4 minutes (Fig. 6C) to 1.9 minutes, with the mean not changed.