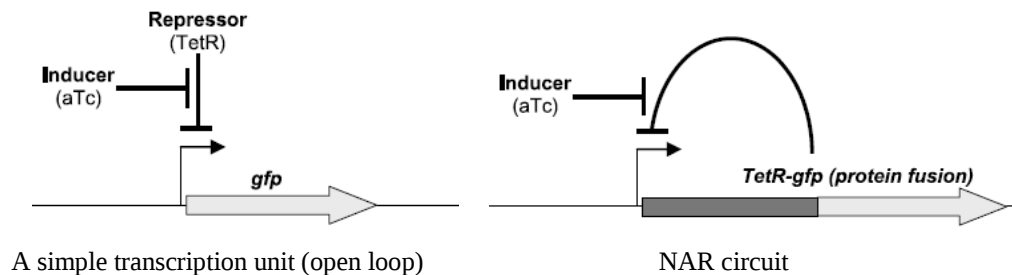


Chapter 4 Network Motifs of Autoregulation (AR) and Feed Forward Loop (FFL)

4.1 Negative AR Motif Speeds up the Response of Transcription Networks (J. Mol. Biol. 323, 785–793 (2002))

Understanding the design principles of gene regulation networks is a major challenge, which requires an effective way to analyze the responsive properties of regulatory motifs. Network motifs are often linked to the rest of the network in a way that preserves their independent dynamical functions. Such a design might allow building complex networks out of circuit elements that can be reliably wired to each other, keeping the proper internal workings of each circuit. It would be important to see whether the full network dynamics can be understood by considering separately the dynamics of each network motif. One common motif is negative autoregulation (NAR), which occurs over 40% of known *Escherichia coli* transcription factors. The effect of NAR on the kinetics of transcription was shown to speed up transcription responses.

Experimental Design: Compare the gene regulation effects of a simple transcription unit and a negative autoregulatory circuit.



The simple transcription unit was represented by cells bearing a plasmid carrying a reporter gene (the green fluorescent protein gene *gfp*) controlled by the *tet* promoter, which is repressed by a constitutively produced repressor *TetR*. When inducer anhydrotetracycline (aTc) appears in the growth medium, expressed *TetR* will bind to the inducer with an extremely high affinity and is inactivated.

To measure the effect of negative autoregulation on the transcription kinetics, a transcription circuit employs a transcriptional factor *TetR*–GFP, which then represses its own production.

Principle of the Experimental Design: The rate of change of the concentration of the gene product $x(t)$ can be described by

$$dx(t)/dt = G(t) - \alpha \cdot x(t)$$

with a generation rate $G(t)$ and an effective decay constant α . Due to the short mRNA lifetime (compared to that of the proteins), mRNA concentration is fixed at a quasi-steady state, which is proportional to $G(t)$.

A simple transcription unit has a constant generation rate $G_1(t) = \beta_1$, which yields a steady-state

concentration of $x_1^{st} = \beta_1/\alpha$. The kinetics of step induction is $x_1(t)/x_1^{st} = 1 - e^{-\alpha t}$. The deviation of $x(t)$ from its steady-state value of a simple transcription unit drops by half each cell-cycle (τ), yielding a rise-time of one cell-cycle $T_{1/2} = \tau$.

Assuming a Michaelis–Menten-like form for the activity of the promoter used, a negative autoregulation circuit has a rate of production of the gene product $x_2(t)$ by $G_2(t) = \beta_2/[1 + (x_2(t)/k)]$. Here k is the dissociation constant of the repressor to its own promoter. The steady-state of $x_2(t)$ becomes

$$x_2^{st} = \frac{\sqrt{k^2 + 4k\beta_2/\alpha} - k}{2} \xrightarrow{\beta_2/\alpha \gg k} \sqrt{k\beta_2/\alpha}.$$

The kinetics approach a simple limiting form $x_2(t)/x_2^{st} \xrightarrow{\beta_2/\alpha \gg k} \sqrt{1 - e^{-2\alpha t}}$ with a rise-time of

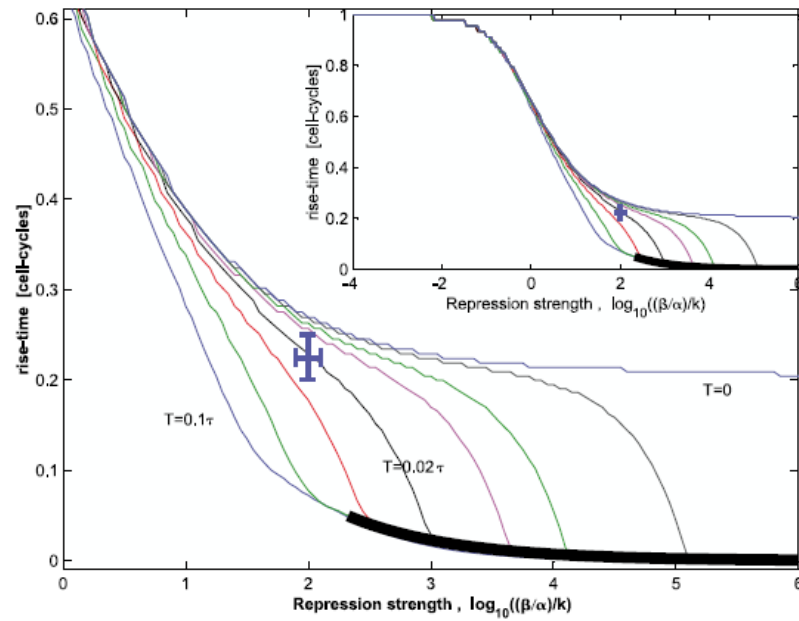
$T_{1/2} = 0.21\tau$. The parameters of the two designs can reach an equal steady-state ($x_1^{st} = x_2^{st}$) by

assigning a relatively weak promoter to the simple circuit and a strong promoter to the autorepressed circuit.

A) Effects of cooperativity: If multiple transcriptional factors are needed on the promoter to

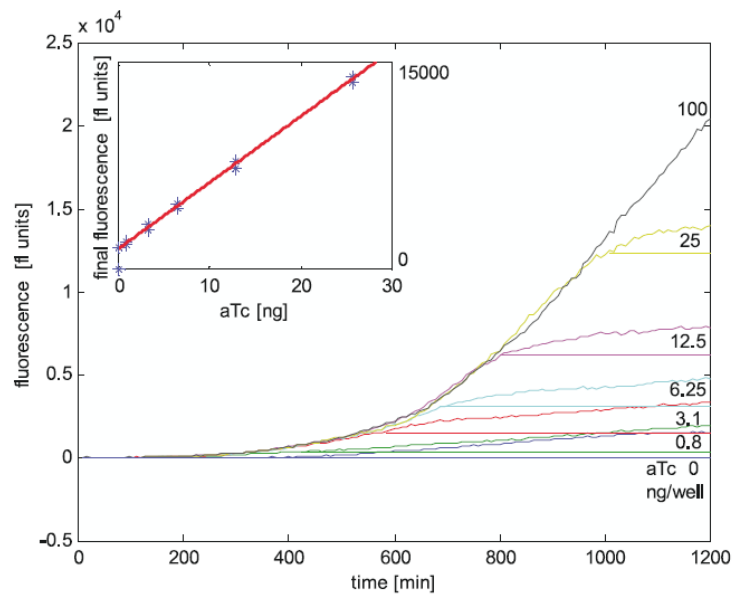
generate transcription $G_3(t) = \beta_3/[1 + (x_3(t)/k_3)^n]$, the lower limit of the rising time $T_{1/2}/\tau$ decreases as the cooperativity n increases.

B) Effects of delays in the formation of proteins: When the delay is not negligible, $x(t)$ in $G(t)$ shall be replaced by $x(t - T)$, which causes a significant effect only when the promoters are so strong that the production of repressor during the delay time T is of the order of the steady state level, that is $\beta T \sim \sqrt{k\beta/\alpha}$. In this case, by the time the first repressors become active, many repressors are already in production. Therefore, the feedback is unable to stop production and a large overshoot in protein concentration can occur (in the thick black line region).



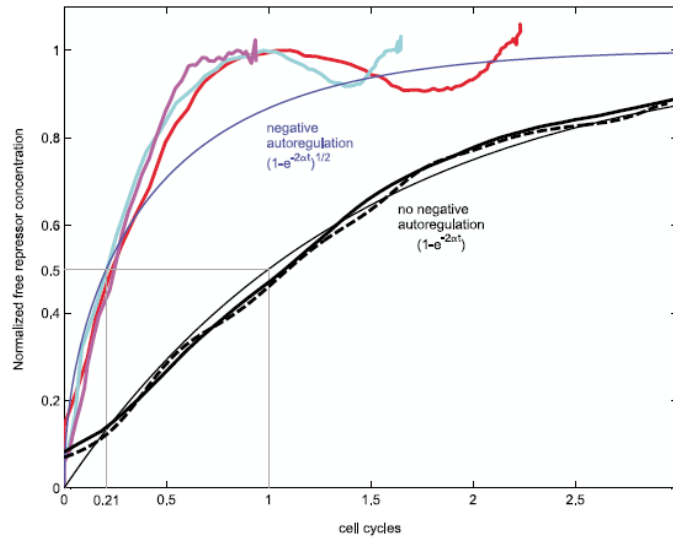
Results:

- A) Titration of aTc: For a fixed amount of aTc, the inducer can be titrated out of the medium by *TetR*. During growth of the cells, *TetR*–GFP fusion protein increases until the concentration of *TetR*–GFP equals that of aTc. From this point on, the NAR kinetics occurs.



Fluorescence (continuous lines) of the negative autoregulatory circuit in response to different concentrations of aTc shows two distinct regimes, an exponential increase in fluorescence followed by a transition to a slower rate of increase.

- B) Experimental Kinetics: Cells from overnight cultures were diluted into fresh medium containing the inducer aTc, which binds and inactivates the repressor *TetR*. After a short lag the fluorescence per cell kinetics agrees with equation.



To observe the effects of NAR on the induction kinetics, one needs to turn on the production of repressor from a low initial concentration of active repressor by using the aTc titration technique as shown in A). The rise-time of the NAR is much smaller than the rise time of an unregulated unit.

Conclusion: Non-self-regulated units have a rise-time of one cell-cycle. Negative autoregulation feedback can reduce the rise time to about one fifth of a cell-cycle.

4.2 Enhanced Stability in Gene Networks with NAR Motif (Nature 405, 590–593 (2000))

It is crucial for a living cell to withstand random perturbations of biochemical parameters to maintain homeostasis and an improved stability in gene regulation networks. The regulation can be provided by NAR motif.

Experimental Design



To construct the autoregulatory system, the tetracycline repressor (*TetR*) was fused to the green fluorescent protein (EGFP) (*TetR-EGFP*) and placed downstream of the lambda promoter containing two tetracycline operators. As controls, the unregulated counterparts were obtained by mutating the *tetR* gene (*TetRY42A*) to eliminate the feedback.

Assuming R =concentration of mRNA, and P =concentration of RNA Polymerase,

$$\frac{dR_{unr}}{dt} = \left[\frac{k_l k_p P}{1 + k_p P} \right] n a - k_{deg} R_{unr} = f(R_{unr}) = f(R_{unr}^*) + \left. \frac{df}{dR} \right|_{R_{unr}^*} \cdot \Delta R_{unr} + O(\Delta R_{unr}^2)$$

$$\frac{dR_{AR}}{dt} = \left[\frac{k_i k_p P}{1 + k_p P + k_r R_{AR}} \right] n a - k_{deg} R_{AR} = g(R_{AR}) = g(R_{AR}^*) + \left. \frac{dg}{dR} \right|_{R_{AR}^*} \cdot \Delta R_{AR} + O(\Delta R_{AR}^2) .$$

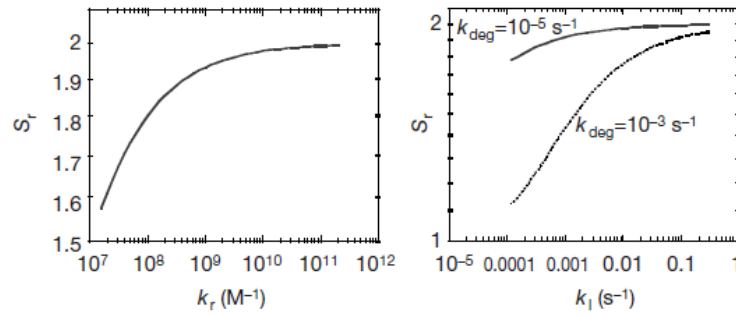
Here k_p and k_r are the binding constants of the RNA polymerase and the repressor, respectively. k_i denotes the promoter isomerization rate from closed to initiating complex, n is the gene copy number, a is the proportionality constant between the mRNA and protein concentration. k_{deg} is the degradation rate of the repressor.

To quantitatively compare the stability of the two systems, a linear stability analysis was performed using the differential equation models of gene circuits. The stability (S) can be determined by linearizing the equations around the steady state, giving $S_r = S_{AR}/S_{unr}$ with

$$S_{unr} = \left. \frac{df}{dR} \right|_{R_{unr}^*} = -k_{deg} \quad \text{and}$$

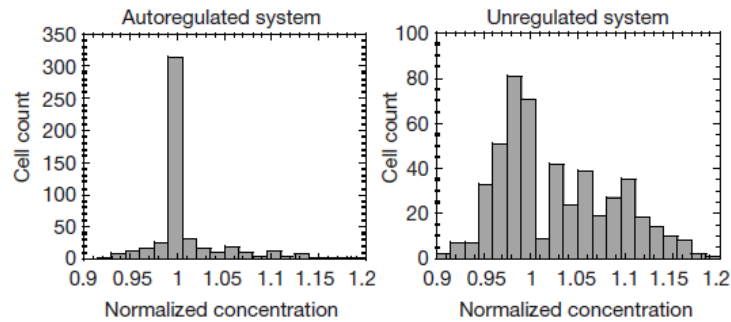
$$S_{AR} = \left. \frac{dg}{dR} \right|_{R_{AR}^*} = -\frac{k_i k_p P}{(1 + k_p P + k_r R)^2} k_r n a - k_{deg} .$$

To calculate the value of relative stability, the parameters of the system are chosen to be $P = 100$ nM, $k_p = 1.5 \times 10^{10} \text{ M}^{-1}$, $k_i = 0.3 \text{ nM s}^{-1}$, $n = 3$, $a = 3.3$, $k_r = 2 \times 10^{11} \text{ M}^{-1}$, $k_{deg} = 10^{-5} \text{ s}^{-1}$. Only one parameter is selected for independent variable: k_r for the left figure and k_i for the right figure.



It is apparent that for all positive values of parameters and steady state concentrations the stability is higher in the autoregulatory system. The autoregulatory system shows a twofold increase in stability over the unregulated one.

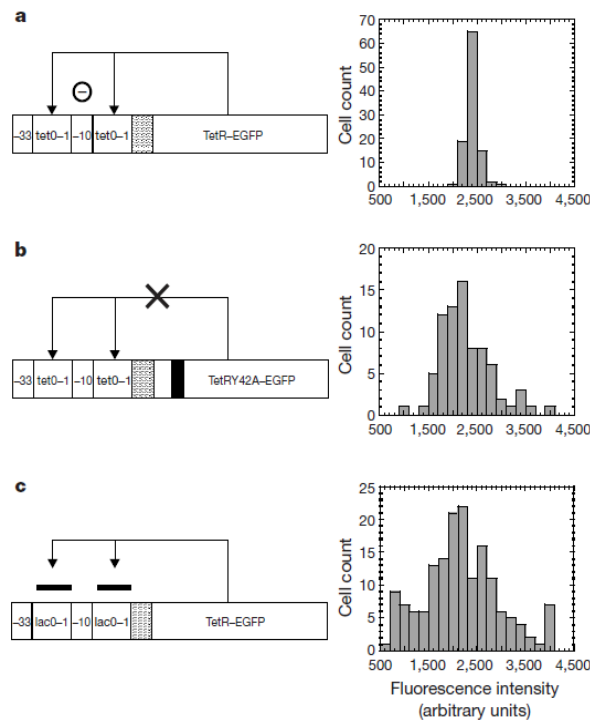
Random perturbations were applied to the steady state of the two systems, which resulted in a variation of the concentration of the transcription factor around the steady state. The distribution is narrower in the autoregulatory system owing to the higher stability.



Results:

By using fluorescence microscopy in the autoregulatory and three unregulated systems, the

expression of the *TetR*–EGFP in the autoregulatory loop exhibits a low steady-state level and high degree of homogeneity with a coefficient of variation (V_C) of 6–9% (a). Decreasing the affinity of the repressor by a mutation increases the variability in the second system and the steady-state concentration of the fusion protein (b). As a third model for an unregulated system, the *tet* operator was replaced by the *lac* operator, so only a negligible nonspecific protein–DNA interaction remains. The expression of *TetR*–EGFP was induced by saturating isopropyl- β -D-thiogalactopyranoside (IPTG) concentration (1mM) to minimize stochastic induction. The distribution of the fluorescence intensities was broader than in any other construct (c).



Conclusion: Roughly 40% of known transcription factors in *E. coli* negatively autoregulate themselves. Negative feedback provides a mechanism to ensure a homogeneous distribution of a transcriptional repressor within optimal concentration limits.

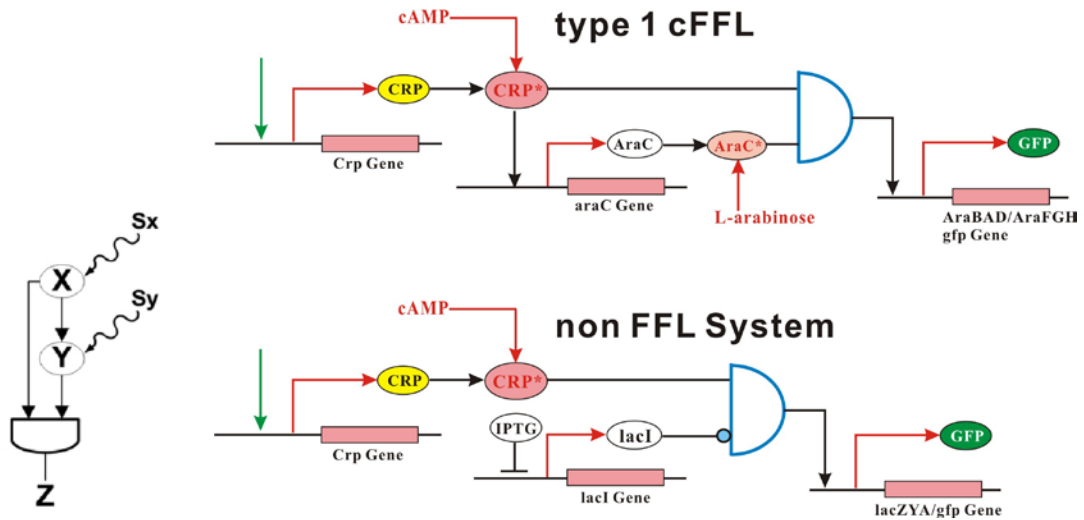
4.3 Coherent FFL (cFFL) Motif Serves as a Sign Sensitive Delay Element (J. Mol. Biol. 334, 197–204 (2003))

One of the most significant network motifs in TRNs is the feed-forward loop (FFL). This motif was first found in *E. coli* and then in diverse organisms.

Experimental Design: In the FFL, transcription factor *X* activates a second transcription factor *Y*, and both activate the output gene *Z*. There are eight types of FFLs, characterized by the signs of the transcription interactions (either *repression* or *activation*). Four of these configurations are termed coherent with the sign of the direct regulation path from *X* to *Z* the same as the overall sign of the indirect regulation path from *X* through *Y* to *Z*. The other four structures are termed incoherent with the signs of the direct and indirect regulation paths *opposite*.

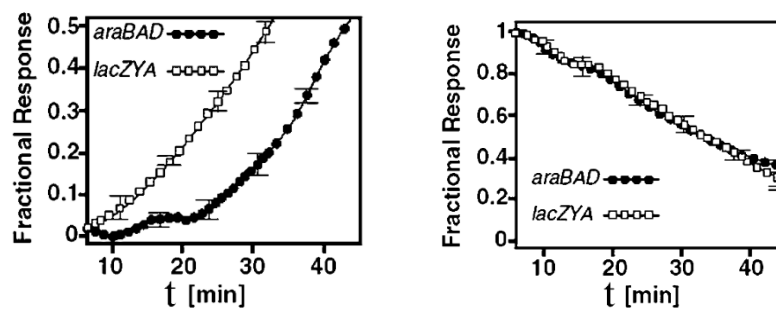
The *ara* system of *E. coli*, which includes the catabolism operon (a set of genes sharing the same mRNA) *araBAD*, and transporters such as *araFGH*, is used to reveal the property of FFL motif

experimentally. Both *araBAD* and *araFGH* are regulated transcriptionally by two transcription factors, *AraC* and *CRP*. *AraC* acts as a transcriptional activator when it binds *L*-arabinose, and as a repressor in its absence. *CRP* acts as an activator when it binds cAMP, which is produced within the cell upon glucose starvation. *CRP* binds the *araC* promoter and enhances the transcription of *araC*. Therefore, in the presence of *L*-arabinose, the *ara* system has the connectivity of a type-1 coherent FFL.



A non-FFL system with the same input S_x is chosen as a control system (the lactose (*lac*) utilization system), in which *CRP* and *LacI* jointly regulate the *lacZYA* operon, but with no transcription regulation of *LacI* by *CRP*. The activity of the promoters is reported by the expression level of green fluorescent protein (*gfp*) gene.

Results: The expression in the presence of both inducers (cAMP and *L*-arabinose/IPTG) is at least an order of magnitude greater than the expression measured when *either* or *both* inducers are missing, indicating that these promoters behave as an AND-gate.



The temporal responses of the *ara* system to cAMP steps were measured by adding saturating cAMP to cells growing exponentially on glucose minimal medium (left). A cAMP OFF step was generated by adding saturating glucose to cells growing exponentially in glycerol minimal medium (right). A significant and reproducible delayed response to cAMP ON steps for *araBAD* relative to *lacZYA* was found to be about 13 minutes at 30°C (left). The delay time in response to ON steps is determined by the time it takes for *Y* to reach levels sufficient to activate *Z*.

In contrast, the response to OFF steps was indistinguishable for all promoters, consisting of simple exponential decay with equal timing (right). The asymmetric behavior of *araBAD* and *araFGH*, with

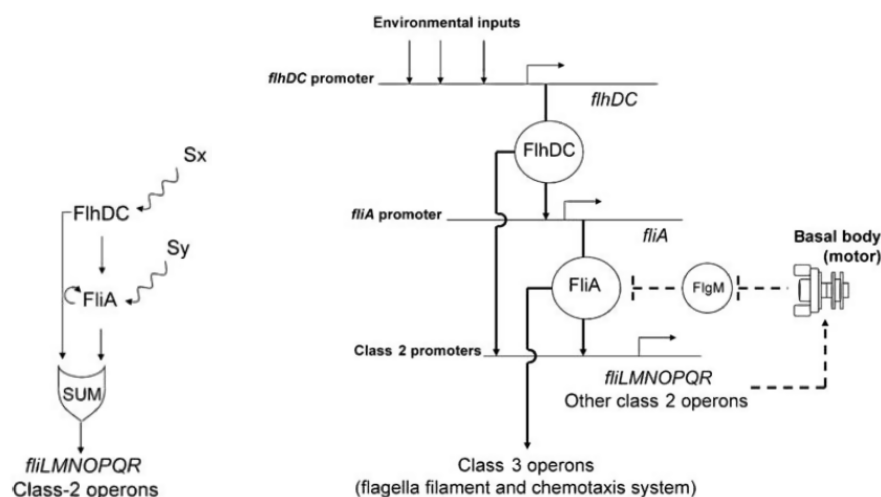
delayed responses to cAMP ON steps but not to OFF steps, is the hallmark of **sign-sensitive delay**.

Conclusion: Bacterium may have an advantage in a rapidly varying environment if it responds asymmetrically when signals turn ON versus OFF. The FFL can protect the target genes from transient cAMP ON signals, allowing them to respond only to persistent stimuli.

4.4 cFFL Motif with a SUM Input Prolongs Flagella Expression in *E. coli* (ref: Molecular Systems Biology 1, 1-6 (2005); doi:10.1038/msb4100010)

In order to understand the function of the FFL, one needs to specify the input function that integrates the effects of X and Y on gene Z . A coherent FFL with an OR gate can carry out an information-processing function of sign-sensitive delay: The output Z responds rapidly when the level of X increases, whereas Z responds only at a delay once X levels decrease. Thus, this gene circuit can protect against transient deactivation. A simple way to implement OR input function is to provide a gene with two different promoters, each responding to one of the inputs. However, the FFL motif deals with the interactions of three genes in isolation. In reality, this circuit is embedded in a network of interactions. It is therefore crucial to experimentally test the dynamical behavior of this motif in living cells.

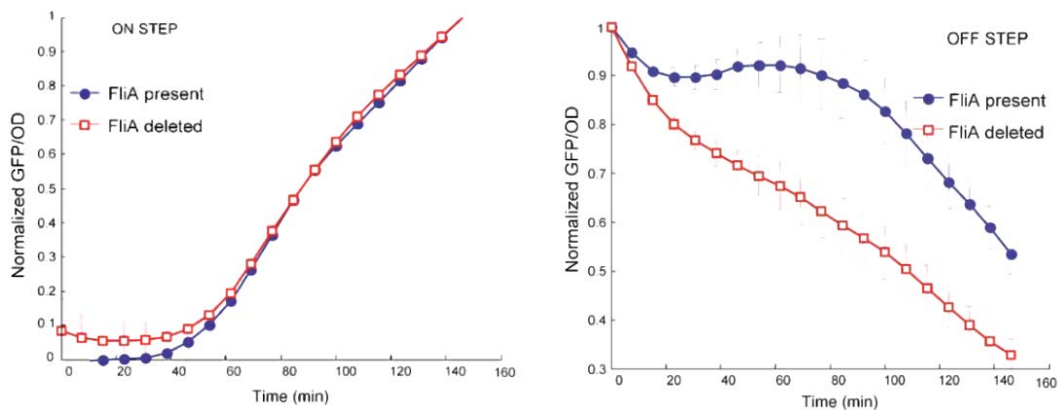
Experimental Design: When growth conditions become mildly unfavorable, *E. coli* produces several rotating flagella and swims away. The flagella biosynthesis network is regulated by a SUM-FFT motif. The master flagella activator X (**FlhDC**) activates a second activator Y (**FliA**). The activators X and Y function additively to activate the genes Z that build the flagella motor (Z represents the flagella class 2 genes arranged in operons such as *fliLMNOPQR*, termed **fliL**). The input S_x is a stimulus that activates X . The input S_y regulates the activity of Y . Here Y also positively regulates its own production to slow the reduction in Y levels following deactivation of X .



Results: The production of *FlhDC* (X) can be turned ON or OFF by means of a chemical inducer *L*-arabinose added externally to the cells. The rate of *FliL* (Z) production was monitored in real time by means of a green-fluorescent protein gene fused to a copy of the DNA regulatory region of the *fliL*.

promoter. As a control, we compared the dynamics to cells in which the gene for *fliA* (Y) was deleted.

To study turn-ON of gene expression, we added an inducer to the cells to initiate the production of X. We found that Z shows rapid production following an ON step of X production. To study turn-OFF of gene expression, we shifted cells growing with inducer for 3 h to a medium without inducer (and with saturating anti inducer D-fucose). We find that the deactivation of Z occurred at a delay of about 60–80 min compared to a cell in which Y is deleted. Thus, the SUM-FFL displays a sign-sensitive delay, with a delay following OFF but not ON steps of X production.

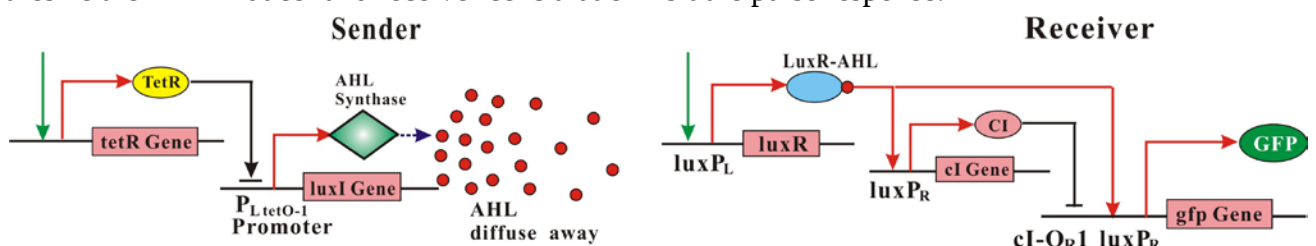


Conclusion: The SUM-FFL can generate a delay in the turn-OFF dynamics of the system, a delay that is dependent on the presence of Y and makes the flagella system insensitive to brief periods in which X is deactivated. It allows the flagella system to turn-OFF only when the proper conditions are sensed for a lengthy period of time.

4.5 Incoherent FFL (iFFL) Produces a Pulse-Generating Gene Expression (PNAS 101, 6355–6360 (2004))

Pulse behavior of TRN is prevalent in many biochemical processes in cells and important in naturally occurring systems. But their operating principles are not well understood quantitatively. Building and studying synthetic networks that exhibit similar behavior can be helpful for an improved understanding of the principles, and for engineering cellular systems for synthetic biology. These efforts attempt to modify the behavior of individual cells to exhibit a desired response. It is also important to observe coordinated behavior in multicellular environments.

Experimental Design: The following figure depicts the genetic circuits for sender cells that synthesize the AHL inducer and receiver cells that exhibit the pulse response.

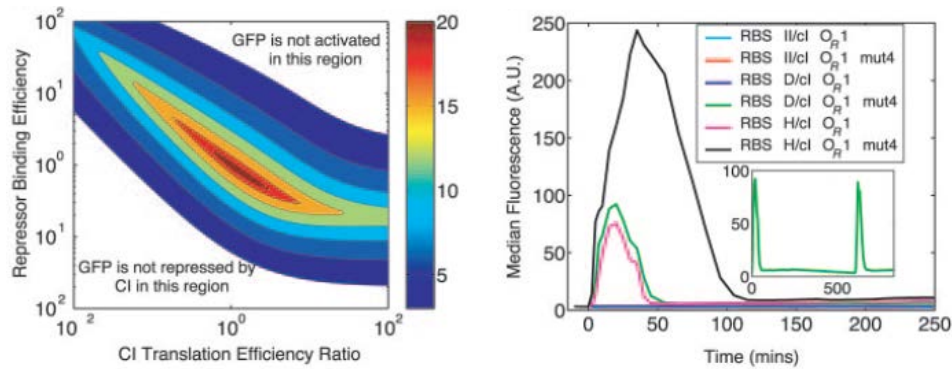


The *LuxI* synthase, which catalyzes the production of AHL, is expressed under the control of the $P_{LtetO-1}$ promoter. The produced AHL molecules diffuse freely from the senders to the receivers. The pulse-generator circuit in a receiver comprises a *LuxR* protein (controlled by the $luxP_L$ promoter), CI

(controlled by $luxP_R$ promoter), and a green fluorescent protein reporter GFP (controlled by $luxP_R$ cI-OR1).

Transcription of both CI and GFP is activated by the *LuxR*-AHL dimer binding the $luxP_R$ promoter. Once CI accumulates in sufficiently high concentrations, it binds the hybrid $luxP_R$ promoter and inhibits further production of GFP. The competition between GFP expression and CI build-up results in transient GFP expression in response to a long-lasting increase in AHL concentration.

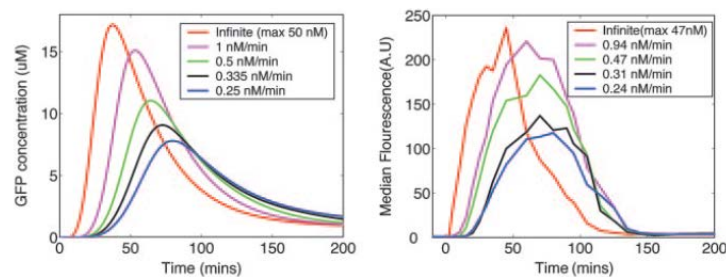
Results: Key characteristics of a pulse include rise time, fall time, width, amplitude, and gain (defined by the difference between the maximum and the final steady state GFP values, divided by the final steady state GFP value).



Simulation on the effect of CI translation efficiency and repressor/operator binding affinity on pulse gain shows that with a high CI translation efficiency and operator binding affinity, GFP levels never rise. In this case, even without AHL, a small leaky CI expression can completely repress $luxP_{R\text{ cI-OR1}}$. In contrast, low translation efficiency and repressor binding affinity result in high GFP expression with little repression.

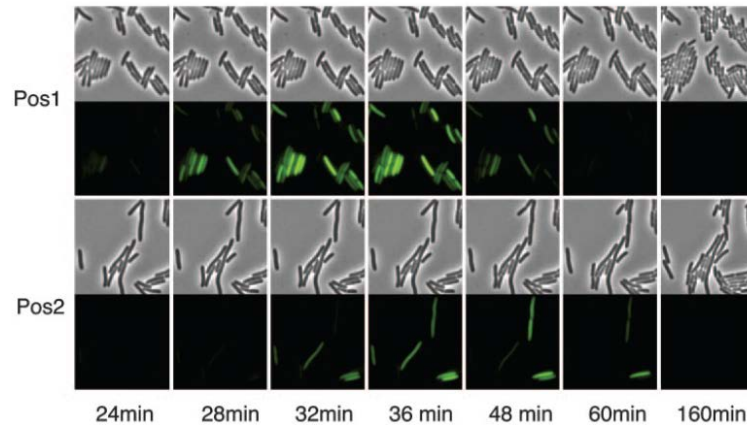
With a proper (weaker than the wild type) ribosome combination site (RBS H) strength and repressor binding strength by mutating the OR1 at single base, it becomes possible to produce a pulse with high fluorescence and extended duration in response to AHL. After AHL induction for 4 h, cells were washed and grown in fresh media without AHL for 6 h. By adding AHL (140 nM) again, the second GFP pulse can be generated with the same intensity levels as the first pulse.

The simulated (left) and experimental (right) responses of the circuit to different rates of AHL increase are presented in the following figures.



When the AHL increase rate is high, the initial buildup of both GFP and CI is high, and therefore, CI quickly shuts down $luxP_{R\text{ cI-OR1}}$ activity. During this window of activity, GFP is produced in large quantities, generating a pulse with short delay and high amplitude. In contrast, when the AHL increase rate is lower, the initial buildup of both GFP and CI is correspondingly lower. It therefore takes longer for CI to shut down $luxP_{R\text{ cI-OR1}}$.

As illustrated above, the receiver cells can differentiate between communication from nearby and far-away sender cells. The following figures show the time-lapsed photographs of circuit behavior at two positions on agar slides. On average, cells that were closer to the senders began fluorescing earlier and displayed a pulse with a higher intensity than cells further away, reflecting the liquid-phase experimental results for different rates of AHL increase.

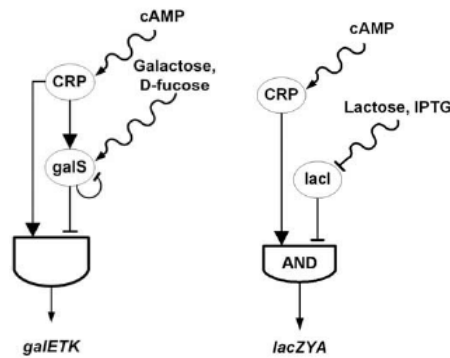


Conclusion: The pulse generating motif based on type-1 incoherent FFL yields a rate-sensing capability. Gradual increases in signaling molecule concentrations are more likely to occur in natural systems. The pulse generator can serve as a model system to understand similar transient and spatiotemporal behaviors found in nature.

4.6 iFFL Accelerates the Response Time of the Gal System in *E. coli* (ref: J. Mol. Biol. 356, 1073–1081 (2006))

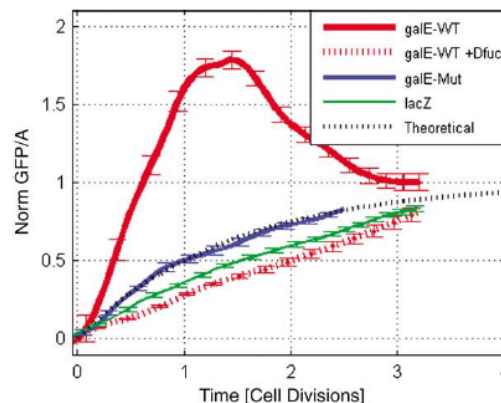
There are two popular FFL types, which are termed the coherent type-1 FFL (C1-FFL) and the incoherent type-1 FFL (I1-FFL). In the I1-FFL, *X* activates *Y* and *Z* while *Y* represses *Z*, and can be employed to accelerate the transcriptional response.

Experimental Design: The dynamics of the I1-FFL in living cells using the *crp-galS-galE* system of *E. coli*. The *gal* system allows *E. coli* to grow on the galactose. Expression of the *gal* genes is inhibited in the presence of glucose. The galactose utilization operon *galETK*, called *galE*, is transcriptionally regulated by *CRP*, an activator induced by glucose starvation (which produces cAMP). The *galE* promoter is also repressed by *GalS*. *GalS* unbinds from the *galE* promoter in the presence of the inducer β -D-galactose (or D-fucose). The gene that encodes the repressor *GalS* is itself positively regulated by *CRP*, so that an I1-FFL is formed. Dynamical expression measurements of the *gal* system in *E. coli* are compared to the control system that does not display the FFL connectivity.



Results: The *galE* promoter activity is enhanced upon depletion of glucose from the medium, resulting in an increase of expression. The normalized dynamics of the *galE* promoter without inducer shows an overshoot, and is accelerated. It reaches 50% of its steady-state level after 0.34 cell generations.

In contrast, the dynamics with saturating D-fucose induction is a monotonic increase that resembles the theoretical solution of constant production and dilution, reaching 50% of the steady-state level after more than one cell generation time. Similarly, the *lacZ* promoter does not show accelerated dynamics by reaching 50% of its steady-state level after more than one cell generation time. The dynamics of a mutant *galE*, in which the main binding site of *galS/galR* was deleted. This mutated promoter loses its responsiveness to D-fucose. The resulting normalized dynamics of this promoter reveals no acceleration following the depletion of glucose.



Response acceleration by the I1-FFL is due to the fact that at early times, *CRP* strongly activates the *galE* promoter, resulting in rapid production. In parallel, *CRP* activates *GalS* production. Thus, at a delay, *GalS* builds up to repress the promoter, locking the system at the desired steady-state promoter activity. The stronger the repression of *Z* (*Gale*) by *Y* (*GalS*), the faster the response time.

Conclusion: The I1-FFL helps to accelerate the metabolic process, and potentially to allow the *gal* system to reach functional protein levels faster, and thus to be ready to use galactose earlier if it appears.