

# Supporting Information

## 1. Experimental supports

### Comparison the fusion construct with the bi-cistronic construct

We constructed the fusion *cI::gfp(aav)* gene in order to have one to one correspondence between the number of CI and GFP(AAV). We confirm that this genetic fusion does not prevent the activator function by referring to the bi-cistronic construct at the dynamics level. If the dynamics of the bi-cistronic construct is identical to that of the fusion construct, these constructs are considered as the same positive feedback system. Based on this assumption, the bi-cistronic construct  $P_{RM} - lacO - rbsA - cI - rbsA - gfp(aav) - T0T1$  was constructed. The  $P_{RM} - lacO$  promoter was positioned upstream of the *cI* gene as the first cistron. Further, the *gfp(aav)* gene was positioned as the second cistron downstream of the  $P_{RM} - lacO$  promoter, ensuring that transcription from  $P_{RM}$  results in the expression of CI and GFP. We induced both of the  $P_{RM}$  wild-type positive feedback systems ( $P_{RM}$  wt-PFS) by 1mM IPTG. We find that the dynamics of the bi-cistronic positive feedback system agrees well with that of the fusion construct (Figure S1). Furthermore, we constructed  $P_{RM} - lacO - rbsA - gfp(aav) - T0T1$  as *cI*-deleted control system. This *cI*-deleted construct did not show the remarkable fluorescence enough to distinguish it from background fluorescence. This is the evidence that the  $P_{RM}$  promoter needs CI to enhance its own activity (Materials and Methods). According to these results, we conclude that the fusion construct did not interfere with the activator function of CI proteins.

### • GFP purification and quantification, the degradation rate of the GFP variant, GFP(AAV)

The plasmid pET27a-GFPmut3\* was made by inserting *gfpmut3\** into BamH1-EcoR1 digested pET28a (Novagen). N-terminal six histidine-tagged GFP was expressed in *E.coli* BL21(DE3) cells and purified by Ni-NTA (QIAGEN) column chromatography. After dialysis and desalination, the protein was concentrated to 6.3 mg/ml and stored at  $-80^{\circ}\text{C}$ . By measuring the fluorescence intensity of GFP solutions at various concentrations, we have found the concentration is proportional to GFP fluorescence. All of the measurements were done in the linear range.

The GFP(AAV) used in this study becomes fluorescent within minutes after transcription. Because GFP(AAV) has *ssrA* tag in the C-terminal, its degradation rate

is not negligible. In order to estimate the half-life of GFP(AAV) *in vivo*, we carry out the following experiment. The cells with the *gfp(aav)* expression plasmid were grown in 1mM IPTG in order to fully produce GFP(AAV). After washing cultures 3 times in PBS, we transferred them to fresh minimal medium without IPTG. Thereafter, 3ml of cultures was sampled, and fluorescence of GFP and  $O.D_{600}$  were measured at various time intervals. It is found that GFP(AAV) is decreased exponentially (Figure S2). The degradation rate is represented as the slopes of the curve. By estimating the slope of the decay, we quantify the half-life of the GFP(AAV) approximately as 60 min. Furthermore, the fact that the stable GFP is not significantly decreased (10% decrease till 180 min) during this measurement meant that the effect of cell division is little.

As shown in Fig. S2, the half time of GFP(AAV) is 60 minutes. The doubling time of *E.coli* is 90 minutes in the experiment. If GFP is stable, the rise-time of no feedback systems is equal to the doubling time of cells. However, if GFP is unstable and cells grow exponentially, the rise-time can be written as ;

$$\gamma = \left( \gamma_{\text{degradation}}^{-1} + \gamma_{\text{dilution}}^{-1} \right)^{-1},$$

where  $\gamma_{\text{degradation}}$  denotes the degradation rate of protein and  $\gamma_{\text{dilution}}$  denotes the dilution rate by cell division [1]. Following this theory and using

$$\tau_{\text{degradation}} = \frac{\ln 2}{\gamma_{\text{degradation}}} \approx 60 \text{ min} \quad \text{and} \quad \tau_{\text{dilution}} = \frac{\ln 2}{\gamma_{\text{dilution}}} \approx 90 \text{ min},$$

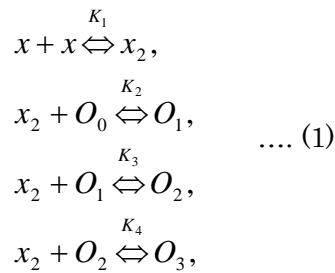
we obtain 36 min as the effective rise-time of no-feedback. However, the rise-time of no-feedback system is 50 min. This difference may be resulted from (1) the initial lag-time in the gene expression systems, *e.g.*, GFP maturation (approximately 5 min), transcription and translation (approximately 5-10min [2]), or (2) over-simplification of the model on protein dilution. We define the effective half-life of protein as 50 min. It leads to  $\gamma = 0.014$ .

## 2. Theoretical supports

- The derivation of the mathematical model of the positive feedback system

We model the  $P_{RM} - lacO$  positive feedback system based on the chemical reactions [2]. The present model is not meant to be a detailed account of the full biochemistry of the system, rather a simple model that captures the essential behavior. The model is the variant of the Collins model. The difference is that we neglect the influence of the cell growth on chemical reactions.

The chemical reactions describing the genetic network are naturally divided into two categories: fast and slow. The fast reactions have rate constants of the order of seconds and are therefore assumed to be in equilibrium compared with the slowly changing variables, which evolve on scales of the order of minutes. Let us denote  $x$ ,  $x_2$ , and  $O_i$  denote the CI monomer, the CI dimer, and DNA promoter sites with  $i$  CI dimers bound, then we may write the fast reactions as

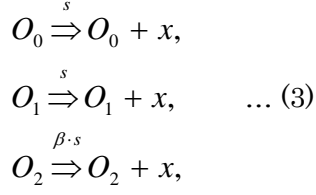


The equilibrium constants for the reactions in Eq.1 are defined by:

$$\begin{aligned}
 K_1 &= \frac{[x_2]}{[x]^2}, \\
 K_2 &= \frac{[O_1]}{[O_0][x_2]}, \\
 K_3 &= \frac{[O_2]}{[O_1][x_2]}, \\
 K_4 &= \frac{[O_3]}{[O_2][x_2]}.
 \end{aligned}
 \quad \dots (2)$$

The slow irreversible reactions are transcription, translation, and protein degradation. Hence, CI monomers are transcribed and degraded at a slow rate. If no repressor is bound to the operator region or a single CI dimer is bound to the first right operator site (OR1), transcription proceeds at a basal rate. If a second CI dimer is bound to OR2, the binding affinity of RNA polymerase to the promoter region is enhanced with a factor of  $\beta$  times, leading to an amplification of transcription. A CI dimer functions

as a transcriptional activator. We write the reactions governing these processes as



where the transcription and the translation are modeled as a single reaction with rate constant  $s$ .

We seek to write a dynamical equation for the slowest process described by the reactions in Eqs.(1)-(3). We neglect the dynamics of CI dimerization because the binding of CI monomers and the dissociation of CI dimers are fast compared to the transcription and the degradation. Thus, although CI monomers are transcribed and degraded at a slow rate, the dynamics of dimerization is not a slowly changing quantity because CI monomers quickly equilibrate with CI dimers. The rates of the transcription proportional to  $[O_i]$  are defined as

$$F(x) = s([O_0] + [O_1] + \beta[O_2]). \quad \dots (4)$$

Utilizing Eqs. (3), we can rewrite Eq. (4) as

$$F(x) = s(1 + K_2[x_2] + \beta \cdot K_2 K_3 [x_2]^2)[O_0]. \quad \dots (5)$$

In this case, the temporal time evolution is given by the rate equation

$$\frac{dX}{dt} = F(x) - \gamma \cdot x. \quad \dots (6)$$

$$X = x + 2x_2 + 2 \cdot [O_1] + 4 \cdot [O_2] + 6 \cdot [O_3], \quad \dots (7)$$

where  $X$  denotes the total CI protein concentration per cell, and  $x$ ,  $x_2$  and  $[O_i]$  are the concentrations of CI monomers, dimers, and dimer-DNA complexes per cell.  $\gamma$  represents the degradation rate of protein and the dilution over the course of one cell-division time.

We can eliminate the number of unbound operator sites  $[O_0]$  by taking the copy number  $[O_T]$  as constant,

$$\begin{aligned} [O_T] &= [O_0] + [O_1] + [O_2] + [O_3] \\ &= (1 + K_2[x_2] + K_2 K_3 [x_2]^2 + K_2 K_3 K_4 [x_2]^3)[O_0]. \quad \dots (8) \end{aligned}$$

This leads to an equation for  $[O_0]$ :

$$[O_0] = \frac{[O_T]}{1 + K_2[x_2] + K_2K_3[x_2]^2 + K_2K_3K_4[x_2]^3} \cdot \dots \quad (9)$$

The production function  $F(x)$  is then obtained by rearrangement as:

$$F(x) = m \frac{s(1 + K_2[x_2] + \beta \cdot K_2K_3[x_2]^2)}{1 + K_2[x_2] + K_2K_3[x_2]^2 + K_2K_3K_4[x_2]^3} \cdot \dots \quad (10)$$

where  $[O_T] = m$ . Now, utilizing Eqs. (6) and (10), we can rewrite Eq. (6) as

$$\frac{dX}{dt} = \alpha \frac{1 + K_1K_2x^2 + \beta \cdot K_1K_2K_3x^4}{1 + K_1K_2x^2 + K_1K_2K_3x^4 + K_1K_2K_3K_4x^6} - \gamma \cdot x, \quad \dots \quad (11)$$

where  $\alpha = m \cdot s$ . The last step is to write  $dX/dt$  in terms of  $dx/dt$ . Utilizing Eqs. (7) and (11), we obtain

$$\begin{aligned} \frac{dX}{dt} = \frac{dx}{dt} & \left[ 1 + 4K_1x + 4K_1K_2x[O_0] + 16K_1K_2K_3x^3[O_0] + 36K_1K_2K_3K_4x^5[O_0] \right] \\ & + \frac{d[O_0]}{dt} \left[ 2K_1K_2x^2 + 4K_1K_2K_3x^4 + 6K_1K_2K_3K_4x^6 \right], \quad \dots \quad (12) \end{aligned}$$

where we have assumed that if the number of monomers is not too small, the number of operator states  $[O_0]$  without bound dimers will be small and rarely changing. We thus can approximate Eq.(12) by letting  $d[O_0]/dt = 0$ . This leads to the governing equation for the evolution of the number of CI monomers,

$$\frac{dx}{dt} = \frac{1}{h(x)} (f(x) - \gamma \cdot x), \quad \dots \quad (13)$$

and

$$f(x) = \alpha \frac{1 + K_1K_2x^2 + \beta \cdot K_1K_2K_3x^4}{1 + K_1K_2x^2 + K_1K_2K_3x^4 + K_1K_2K_3K_4x^6} \cdot \dots \quad (14),$$

$$h(x) = 1 + 4K_1x + 4K_1K_2x \cdot [O_0] + 16K_1K_2K_3x^3 \cdot [O_0] + 36K_1K_2K_3K_4x^5[O_0] \cdot \dots \quad (15)$$

$$[O_0] = \frac{m}{1 + K_1K_2x^2 + K_1K_2K_3x^4 + K_1K_2K_3K_4x^6}, \quad \dots \quad (16)$$

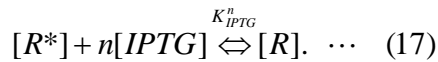
In the positive feedback systems of both  $P_{RM}$  L1L1R3 and  $P_{RM}$  R3R2R3, the mathematical model can be also derived from chemical reactions. In the  $P_{RM}$  L1L1R3 positive feedback system, the promoter has three operator sites, OL1, OL1 and OR3. Comparing with the  $P_{RM}$  wild type positive feedback system, OR1 and OR2 are substituted by OL1, respectively. The binding affinity to OL1 is slightly higher than that of OR1 according to the *in vitro* study. If no repressor is bound to the operator region or if a first CI dimer is bound to the upstream OL1, transcription proceeds at a basal rate. And if a second CI dimer is bound to the other OL1, the rate of the feedback activation enhances by  $\beta$ . In the same way as well as the wild type, we obtain the

model equation given by Eq. (13)-(16).

Next, in the  $P_{RM}$  R3R2R3 positive feedback system, the promoter has three operator sites in order, OR3, OR2 and OR3. OR1 is substituted by OR3. The CI binding affinity of OR3 is lower than that of OR1. CI dimer is hard to bind to the first operator site. In this case, it is assumed that CI dimer preferentially binds to the second operator OR2 [2]. If no repressor is bound to the operator region, or if a first CI dimer is bound to the OR2, transcription proceeds at a basal rate. In turn, if a second CI dimer is bound to the upstream OR3, it functions as the transcriptional activator. The transcription proceeds at  $\beta$  times of the basal rate. On the contrary, if a second CI dimer binds to the downstream OR3, it functions as the transcriptional repressor because the binding competition with RNA polymerase occurs. In this case, the transcription does not proceed. Of course, if the system is on the state in which three operator sites are occupied, the transcription does not occur. These chemical reactions can be described by Eq.(13)-(16).

### The derivation of the mathematical model of the lac system

We give a derivation of the model for the *lacO* system [3]. The transcription factor, the LacI repressor, forms a tetramer and binds to the DNA at a site overlapping the binding site of the  $\sigma^{70}$  unit of the RNAP. Thus it physically suppresses the binding of the RNAP to the DNA, and so transcription is not initialized. LacI is deactivated by IPTG: once IPTG binds to LacI a conformational change is induced in the LacI structure; due to this change the LacI-IPTG complex is unable to bind to the DNA, the  $\sigma^{70}$  site is free to accept the RNAP, and transcription is initialized. We assume that the concentrations of the inducers IPTG are much higher than those of the LacI. Only a small portion of the total amount of the inducers actually binds to the LacI. The corresponding chemical reaction between LacI and IPTG is:



with the following detailed balance equation:

$$[R] = [R^*] \cdot [IPTG]^n \cdot K_{IPTG}^n, \quad \dots \quad (18)$$

$$[R_T] = [R] + [R^*].$$

Solving for  $[R^*]$  we find

$$[R^*] = \frac{[R_T]}{1 + (K_{IPTG}^n \cdot [IPTG]^n)}. \quad \dots \quad (19)$$

The chemical reactions in the DNA-protein level obey the following equation:



where  $[R^*]$  denotes the active LacI repressor, while  $[O]$  and  $[OR^*]$  denote the unoccupied operator region and the repressor-operator complex, respectively. Equilibrium constant for the reaction is given by

$$K_R = \frac{[OR^*]}{[O][R^*]}, \quad \dots \quad (21)$$

$$[O_T] = [O] + [OR^*], \quad \dots \quad (22)$$

where  $[O_T]$  denotes the total operator concentration, and  $[O_T]$  is constant.

Define now an active state as a state in which transcription occurs (that is RNAP has bounded to the promoter and has initiated to transcribe). Accordingly, we define the total promoter activity

$$\alpha = \alpha_1[O] + \alpha_2[OR^*]. \quad \dots \quad (23)$$

$\alpha_2$  is the basal level of transcription under the LacI repression. The fact that  $\alpha_2$  is not equal to zero can be understood as a kind of leakage from the promoter.  $\alpha_1$  denotes the transcription from the single unoccupied operator.

Solving the system of linear equations by Eqs.(19), (21) and(22) and substituting the solution into the definition Eq.(23) yields the final result:

$$\alpha = \alpha_3 \frac{s_I + [IPTG]^n}{K_I + [IPTG]^n}. \quad \dots \quad (24)$$

where the repression factor  $K_I = (1 + K_R[R_T]) / K_{IPTG}^n$  describes how tightly LacI can

regulate gene expression, and  $s_I = \left(1 + \frac{\alpha_2}{\alpha_1} K_R[R_T]\right) / K_{IPTG}^n$  is the factor that describes

the leakiness of the *lac* regulation.  $\alpha_3 = \alpha_1[O_T]$  is the maximal promoter activity. This promoter activity exhibits sigmoidal function of [IPTG]. This sigmoidal behavior arises from the fact that the binding of IPTG to any one of four possible sites on the LacI tetramer is sufficient to interfere with LacI activity. We set  $n = 2$  with reference to experimental evidence [4]. Finally, the promoter activity can be also be expressed generally as:

$$\alpha(I) = \alpha \frac{s_I + [IPTG]^2}{K_I + [IPTG]^2}. \quad \dots \quad (25)$$

In our positive feedback systems, CI dimers as well as LacI complexes bind to their specific palindromic operator within  $P_{RM}$ -*lacO* in a cooperative manner, respectively. In the absence of the inducer IPTG, CI dimers and LacI complexes bind to their own operators. In such a situation, the LacI complex dominates over promoter function. On the other hand, In the presence of IPTG, LacI complexes cannot bind to the operator. In this situation, CI dimers dominate over promoter function. Therefore, we assume the  $P_{RM}$ -*lacO* activity is given as

$$\alpha = \alpha_1 \frac{s_I + [IPTG]^2}{K_I + [IPTG]^2} \cdot \frac{1 + K_1 K_2 x^2 + \beta K_1 K_2 K_3 x^4}{1 + K_1 K_2 x^2 + K_1 K_2 K_3 x^4 + K_1 K_2 K_3 K_4 x^6}. \quad \dots \quad (26)$$

This is a gene switch with the AND function. This model can represent the our genetic switches (Figure. S3)

#### • The improvement of the agreement between experiment and simulation

The time-evolution of gene expression in the experiment shows the gradual and relatively faster increase in the initial phase than the simulation. We consider these differences as the averaging effect that is resulted from averaging out the variability of individual cells. The cells have fluctuations of the protein concentration. Stochastic fluctuations become dominant when the protein concentrations are low. If the initial conditions differ, the rise-times become different values. Further, because we observe the GFP fluorescence as the mean-value from the population of cells, the kinetics of gene expression can be averaged out. Under this consideration, we simulated the dynamics of 1000 cells on the assumption that the initial condition distributes obeying the normal distribution. The dynamics with the fluctuation agrees better than that without fluctuations (Supporting information, Figure S4). The gradual increase at the early time may be derived from the cell-cell variability, although they must be confirmed by the experiment at the single cell level.

We also find that the bad fits for the later stages in the  $P_{RM}$  R3R2R3. Though we have obtained  $\alpha = 71$  by the random parameterization, we can make a better agreement with the experimental result by increasing  $\alpha$  ( $\alpha \approx 76$ ) (Figure S4). However, the increase of  $\alpha$  by 10 % is the remarkable because the standard deviation of  $\alpha$  is 5% or so. It means that the maximal promoter activity of the  $P_{RM}$  R3R2R3-PFS is larger than that of the other  $P_{RM}$  PFS. The point mutation may cause the unexpected change. As possible mechanisms, we speculate that (1) the binding affinity of RNA polymerase to the  $P_{RM}$  promoter is increased, and that (2) The transcriptional complex of RNA polymerase becomes easier to initiate transcription.

## References

- [1] Rosenfeld, N., Elowitz, M.B., & Alon, U. (2002). Negative autoregulation speeds the response times of transcription networks, *J. Mol. Biol.*, **323**, 785-793
- [2] Isaacs, J., Hasty, J., Cantor, C.R. & Collins, J.J. (2003). Prediction and measurement of an autoregulatory genetic module. *Proc. Natl. Acad. Sci. USA*, **100**, 7714-7719.
- [2] Little, J.W., Shepley, D.P., & Wert, D. (1999). Robustness of a gene regulatory circuit. *EMBO.J.*, **18**, 4299-4307.
- [3] Setty, Y., Mayo, A.E., Surrete, M.G. & Alon, U. (2003). Detailed map of a cis-regulatory input function. *Proc. Natl. Acad. Sci. USA*, **100**, 7702-7707.
- [4] Yagil, G. and Yagil, E., (1971). On the relation between effector concentration and the rate of induced enzyme synthesis. *Biophys.J.* **11**, 11-27

## Figure Legends

### Figure S1.

Comparison the fusion construct with the bi-cistronic construct. We induced both of the  $P_{RM}$  wild-type positive feedback systems ( $P_{RM}$  wt-PFS) by 1mM IPTG. Gene expression dynamics of the bi-cistronic  $P_{RM}$  wt-PFS (blue-circle) agrees well with that of the fusion  $P_{RM}$  wt-PFS (green-square). Inset figure shows the normalized dynamics of gene expression.

### Figure S2

Stability of GFP variants in *E.coli*. Gray-circle: JM109 expressing GFP mut3\*, Black-square: JM109 expressing GFP(AAV), Gray-square: JM109 expressing GFP(LVA).

### Figure S3

Promoter activity as a function of IPTG concentration. Blue: promoter activity of the NFS, Red: promoter activity of the  $P_{RM}$  wt-PFS. The blue and red lines represent fits to the data by using Eq.(25) and Eq.(26), respectively.

### Figure S4

Simulated gene expression dynamics with considering the fluctuation of the initial condition. Shown is Normalized total protein concentration over time. Experiments (Points), and the mathematical model using the quantified parameters (Solid line). Blue: the NFS, Red: the  $P_{RM}$  wt-PFS, Green: the  $P_{RM}$  L1L1R3-PFS, Orange: the  $P_{RM}$  R3R2R3-PFS.

Figure S1

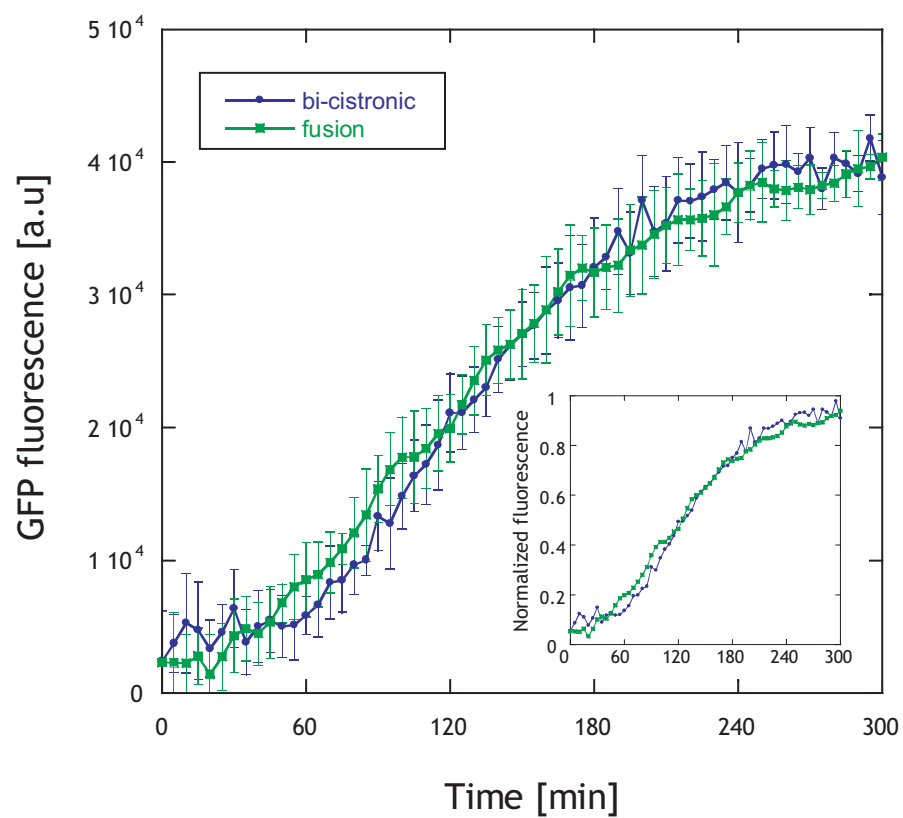


Figure S2

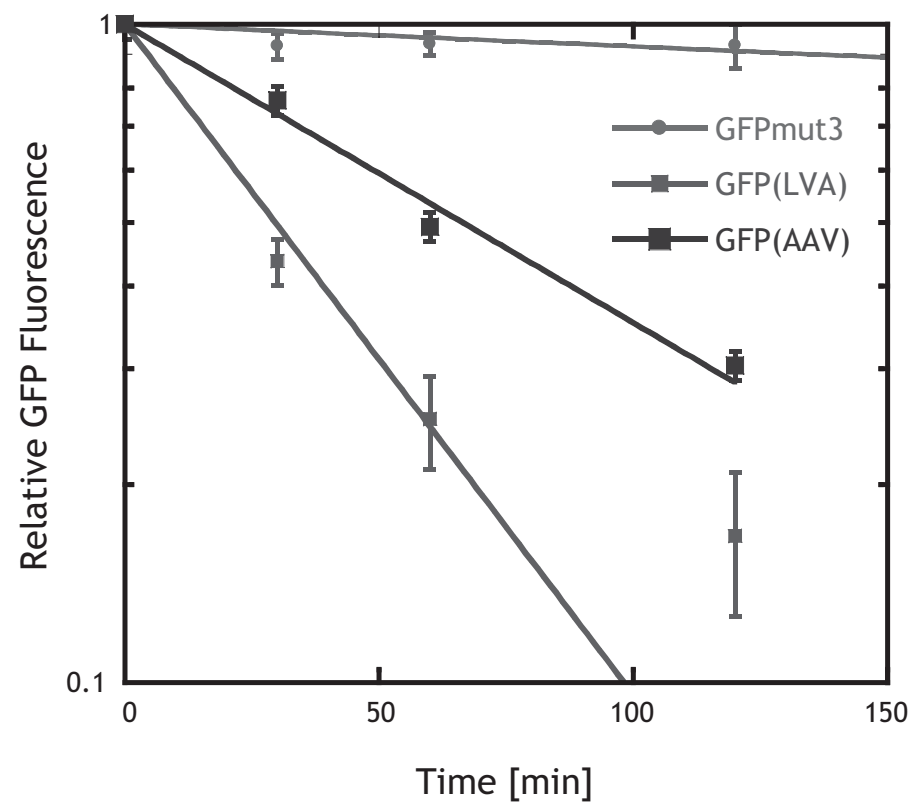


Figure S3

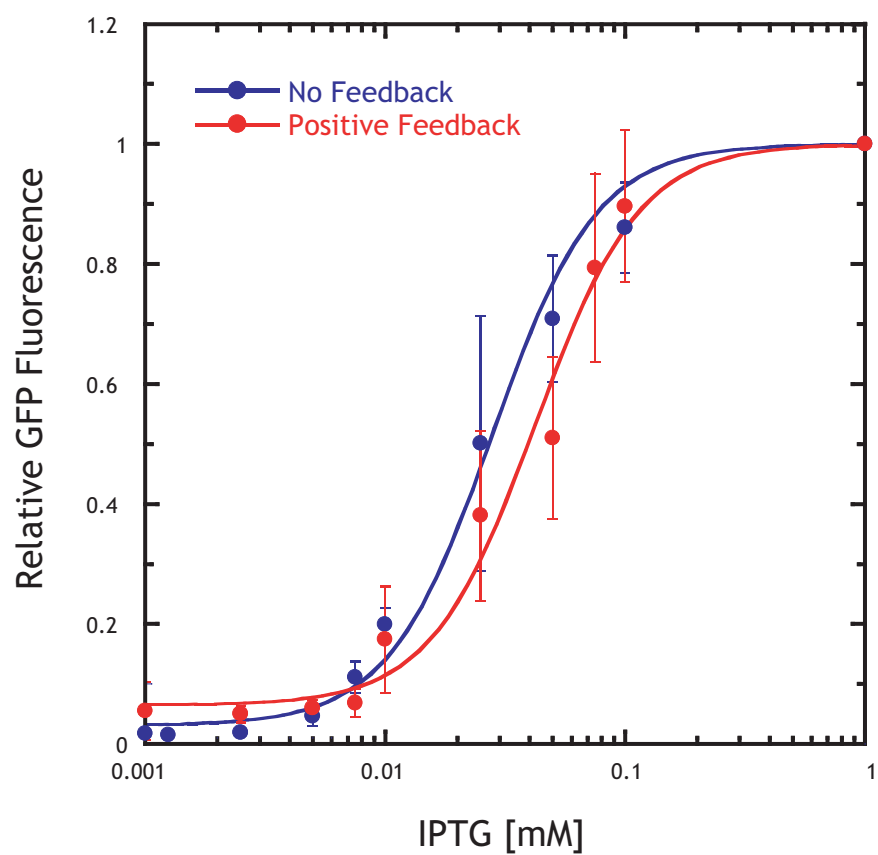


Figure S4

