DESIGN AND IMPLEMENTATION OF THREE INCOHERENT FEED-FORWARD MOTIF BASED BIOLOGICAL CONCENTRATION SENSORS

Robert Entus
Keck Graduate Institute
Claremont, CA 91711, USA.
(robert_entus@kgi.edu)

Brian Aufderheide[2]
Keck Graduate Institute
Claremont, CA 91711, USA.
(brian_aufderheide@kgi.edu)

Herbert M. Sauro[1]
Department of Bioengineering
William H. Foege Building
Box 355061
University of Washington
Seattle, WA 98195, USA.
(hsauro@u.washington.edu)

Word Count: 6560

Abbreviations:

EC	Elongation complex
FFL	feed-forward loop
GFP	Green Fluorescent Protein
IPTG	Isopropyl-β-D-thiogalactopyranoside
LB	Luria Broth
ORF	Open reading frame
PCR	polymerase chain reaction
RBS	Ribosome Binding Site
T7-RNAP	T7 RNA polymerase T7-RNAP
SAM	S-Adenosyl methionine
SBW	Systems Biology Workbench
TE	Translational Efficiency

[1] Corresponding author
[2] Provided material support

ABSTRACT

Synthetic biology is a useful tool to investigate the dynamics of small biological networks and to assess our capacity to predict their behavior from computational models. In this work we report the construction of three different synthetic networks in *Escherichia coli* based upon the incoherent feed-forward loop architecture. The steady state behavior of the networks was investigated experimentally and computationally under different mutational regimes in a population based assay. Our data shows that the three incoherent feed-forward networks, using three different macromolecular inhibitory elements, reproduce the behavior predicted from our computational model. We also demonstrate that specific biological motifs can be designed to generate similar behavior using different components. In addition we show how it is possible to tune the behavior of the networks in a predicable manner by applying suitable mutations to the inhibitory elements.

Summary: This paper describes the construction of three different feed-forward networks. Each network is characterized at steady state and is shown to behave as a concentration detector. A simple computer model is constructed which generates predictions confirmed experimentally by mutational studies on each network.

Keywords: Feed-forward networks, Modules, Simulation, Synthetic Biology

1. Introduction

Cells perform information processing functions using biochemical networks (Bray, 1995; Hartwell et al., 2002; Sauro and Kholodenko, 2004) and there is evidence to suggest that that such networks contain recurring functional patterns or "motifs" (Shen-Orr et al., 2002; Milo et al., 2002; Lee et al., 2002). Network motifs occur far more often than equivalent randomized networks in *Escherichia coli* (Shen-Orr et al., 2002; Milo et al., 2002) and *Saccharomyces cerevisiae* (Milo et al., 2002; Lee et al., 2002).

One of the common motifs found in both *E. coli* and *S. cerevisiae* (Shen-Orr et al., 2002; Milo et al., 2002) is the feed-forward loop (FFL). Fig. 1 illustrates a typical FFL, where there is a common transcription factor for two genes, and the third gene is regulated in a feed-forward fashion. The input, X, modulates the activity of a target gene (Z), directly, and indirectly through the gene product, of another gene (Y). The interaction between X and Y, at the promoter region of Z determines the rate of transcription of Gene Z. This type of architecture has been shown (Mangan et al., 2003) to lead to two types of dynamics depending on the nature of the regulation which occurs at the target gene Z. If X, Y are both activators, then the gene circuit acts as a low pass filter, *i.e.* it is able to filter out transient signals and transcribe only when the input signal is long lived (Mangan et al., 2003). If X, Y regulate Z as an activator and repressor respectively then the system can act as a biological concentration sensor, since the delayed response of X through Y, tends to suppress activity of Z. This has been suggested to be a mechanism for speeding response times in transcriptional networks (Mangan et al., 2003).

Ishihara et al. (2005) showed *in silico* that the concentration sensitive bell shaped pattern described by Basu and colleagues (2005) could be produced by the FFL architecture, using three genes instead of five. We have built three networks that satisfy the three gene feed-forward topology capable of validating modeling predictions. All three networks, described in detail later, are based on the tunable expression of T7 RNA polymerase (T7-RNAP), under the control of the lac repressor. The three networks target different macromolecular interactions: protein:DNA, RNA:RNA, and protein:protein.

Of particular interest to us was the steady state response of the networks, we were not concerned explicitly with time dependent changes. By focusing on the steady state response we were able to justify making measurements on cell populations rather than single cells. In addition we assumed that for a given set of parameters, our synthetic networks only admitted a single unique steady state. This unique steady state would however be a function of the input signal and strength of the feed-forward inhibitory and activating loop. This was a reasonable assumption because our networks did not contain any negative or positive feedback loops which could destabilize the dynamics and result in multiple steady states.

The output from the networks was measured using GFP fluorescence from approximately 10^8 *E. coli* cells. The assay conditions minimize cell division to approximately one division every four to five hours. In order to determine when the networks reached steady state we tracked the rate of change of GFP concentration (dGFP/dt = production - degradation) adjusted for cell number (Fig. 4). Once the rate of change of GFP reached a

low level we assumed that the synthetic networks had settled to their steady state operating points. All subsequent measurements where then made from this point in the experiment. This follows behavior expected from earlier work done by Setty and colleagues (2003) who were using similar experimental conditions.

In this paper we describe in detail the experimental construction of three different FFL networks. The output of each FFL network (in the form of GFP) was measured as a function of the input inducer, Isopropyl-β-D-thiogalactopyranoside (IPTG). In addition each network is subjected to mutations in order to tune their behavior. We also complement the experimental work with a simple computational model of a generic FFL network. We apply the same mutational changes to the model in order to make predictions. These predictions are confirmed by the experimental studies.

2. Methods

2.1 Gene circuit construction

The networks were assembled from *E. coli* K12 genes and promoters, a GFP gene (Biobricks BBa_E0040) and T7-lysozyme (Novagen). The network components were amplified by polymerase chain reaction (PCR) using KOD DNA polymerase from Novagen and an Eppendorf Mastercycler Gradient thermal cycler. Novel restriction sites, for example, BspHI and SacI, were added to the 5' and 3' end respectively for directed cloning. PCR conditions were 95 °C for 1 min, followed by 25 cycles of 95 °C for 30 s, 60 °C for 1 min and 72 °C for 1 min, and a final step of 72 °C for 2 min. The products were gel purified with 0.8% agarose gel and Wizard gel purification kit (Promega). The

individual genes were cloned into a pBR322 derivative and verified. This plasmid backbone contains a T7 promoter, T7 terminator, and *lacI* gene and, kanamycin resistance gene. The genes were shuttled into a pCYC184 linearized with PshAI containing a chloramphenicol resistance gene. To create the *metR* promoter, we used *E.coli* K12 DH5a genomic DNA as a template and cloned the intragenic region between *metE* and *metR* into pCYC184 containing a GFP gene. Point mutations were introduced by gel purified primers (Integrated DNA Technology) containing changes of interest on the 5' end.

2.2 Cell growth and expression experiments

All plasmids were inserted into the *E. coli* strain BL21(DE3) via electroporation. Cells or experimentation were grown overnight in LB media supplemented with 0.45% glucose, 0.15% galactose, and 50 µg/mL Chloramphenicol and/or 50 µg/mL kanamycin antibiotic. Cultures were inoculated 1:300 in LB media containing identical supplements and grown in a 37 °C incubator shaking at 250 r.p.m. for approximately three hours, to an A_{600} nm between 0.6 and 0.8. Cultures were diluted with 3-4 volumes of ddH_2O and centrifuged at 4500 g for 4 min. The pellet was washed with 6 mL ddH_2O and centrifuged at 4500 g for 4 min. The resulting pellet was resuspended in 2x M9 minimal media (2x M9 salts, 0.4% cas amino acids, 2 mM $MgSO_4$, 0.3 % glucose, and 0.1% galactose). 60 µL cells/well were added to 96 well plate (Nunc) containing 60 µL ddH_2O containing varying concentrations of IPTG and 100 µg/mL Chloramphenicol and/or 100 µg/mL kanamycin antibiotic as required. Each well was then covered with 100 µL mineral oil and grown with orbital shaking in a Victor3 multiwell fluorimeter at 27 °C. The

fluorescence (excitation: 395 nm, emission: 535 nm) and the absorbance (600 nm) were measured every 8 min for 12 hrs. Each network was examined once steady state was reached; determined when the rate of change of GFP reached a low value (Fig. 4).

2.3 Simulation methodology

The model of the FFL network was constructed from a set of differential equations. Each differential equation in the model describes the rate of change of protein or RNA depending on the physical system. Solutions to these equations can be generated by any suitable simulator package; we used SBW to generate all the plots (Sauro et al., 2003). SBML and Jarnac scripts are available at our web site (www.sys-bio.org) and will be submitted to biomodels.net.

3. Modeling

In this section we present an intentionally simplified model for the feed-forward motif (Fig. 2). This model is the simplest that can be used to qualitatively describe the three networks presented in this paper. In Fig. 2, we will assume that p1 activates G1 and G2; p2, which is the gene product of G1 acts as a repressor of G2. Based on these assumptions, the following rate laws were derived using the method proposed by Shea-Ackers (Ackers, 1982) for transcriptional binding kinetics.

$$(1) \qquad \frac{d[p_2]}{dt} = \frac{t_1 a_1 [p_1]}{1 + a_1 [p_1]} - \gamma_1 [p_2]$$

$$(2) \qquad \frac{d[p_3]}{dt} = \frac{t_2 b_1 [p_1]}{1 + b_1 [p_1] + b_2 [p_2] + b_3 [p_1][p_2]} - \gamma_2 [p_3]$$

At steady state, before G1 is saturated with p1, we have p2 proportional to p1.
For small values of p1, p3 increases proportionally to p1, whereas at large p1 values, p3 decreases proportionally to 1/p1. The net effect is that p3 reaches a maximum at an intermediate value of p1 creating a tunable module that acts as a band pass filter.

At low input concentration, p1 transcribes G2 and G1, and hence as its input level increases, p2 tends to grow. Binding of p2 at G2 halts further transcription of G2. This module is aptly named an "amplitude filter", originally called as a band detector (Basu et al., 2004; Basu et al., 2005; Kaern et al., 2006), since its output is maximal for a specific range of input. Such biphasic responses have also been discussed in other systems (Wolf and Arkin, 2003; Mayya and Loew, 2005). Recently Ishihara et al. (2005), discussed the band properties of such networks and used this to explain pulsed behavior and patterning in *Drosophila* developmental processes.

3.1 Effects of mutations at G2

We can make mutations at G2 to alter, with varying degrees, the ability of p2 to bind G2. In Fig. 3 we illustrate three computer simulations where we plot the concentration of p3 versus the input concentration of p1. The simulation data have been normalized with respect to the peak high to emphasize the lateral shift in the peak response. Thus as the inhibition on G2 is increased the peak response shifts to the left. It is worth noting that the overall magnitude of the signal is also decreased as the affinity of p2 to G2 increases. The resulting decrease in the steady state level of p3 causes the peak signal to shift to the

left. In practical terms, this means that smaller amounts of p2 are required to achieve the same effect.

Such mutations in G2 could be used to engineer the shape of the amplitude filter, by changing the band position. Combinations of mutations can thus be used to alter the band pattern to suite a specific engineering need. In this paper we illustrate mutations that change the binding affinity and show that it alters the band position as predicted by the model.

4. Results

4.1 Unregulated system

The two forms of the unregulated networks (Fig. 5a,b) were measured for fluorescence and corrected for cell density. In the first network GFP production is driven by T7 RNA polymerase (T7-RNAP) under the control of the lac promoter. Increasing IPTG results in increasing amounts of T7-RNAP being produced (Fig. 5c). Fig. 5a shows the basic network under investigation where the T7 binding site is located ~70 bases upstream of the GFP start sequence.

To create an additional regulated network, the promoter region of the GFP gene was modified to contain additional regulatory elements, the *metR* promoter. This promoter is regulated by MetJ through the binding of "met boxes" to down-regulate transcription and is involved in methionine regulation (Weissbach and Brot, 1991). The *met* operator in *E.coli* consists of tandem repeats of eight base pair sequences, homologous to a palindromic consensus AGACGTCT, known as "met boxes" (Belfaiza et al., 1986).

In Fig. 5b a native *E. coli* promoter, *metR*, has been inserted between the T7 binding site and the protein start sequence resulting in approximately 200 bases between the T7-RNAP binding and protein start. The functionality of the promoter region was verified by transforming an *E.coli* DH5a strain lacking T7-RNAP, and measuring the resulting fluorescence. The cells fluoresced in a dose dependent fashion with increasing methionine, indicating functionality of the met boxes, but were unresponsive to IPTG, as expected. When transformed into BL21(DE3) the T7promotor-*metR-gfp* behaves in an equivalent fashion to the T7promotor–*gfp* system in the range of IPTG that we are investigating in this paper.

Gene translation can be influenced by the upstream region the gene. As such, it is unlikely that the efficiency of GFP production between the two systems is identical. We address this issue by scaling the output to an internal standard for each construct. The unregulated networks show similar response to low concentrations of IPTG that becomes saturated at higher concentrations (Fig. 5c). Both unregulated responses show a sigmoidal response with Hill coefficients of 1.93 and 1.42 respectively. The *metR* promoter system has both positive and negative regulators. In this work we have focused on MetJ which is known to effectively repress even in the presence of activators.

4.2 MetJ:*metR promoter Network*

In order to produce the FFL topology, the *metR* promoter was inserted between a T7-RNAP binding site and a GFP gene encoded on the pCYC184 derivative. With the introduction of the *metJ* gene on a pBR322 derivative under the control of T7-RNAP,

GFP is produced from a T7-RNAP/*metR* hybrid promoter region (Fig. 6a) that contains binding sites for both a strong inducer, T7-RNAP, and repressor, MetJ.

This promoter is regulated by MetJ through the binding of "met boxes" to down-regulate transcription. The *met* operator in *E.coli* consists of tandem repeats of eight base pair sequences, homologous to a palindromic consensus AGACGTCT, known as "met boxes" (Belfaiza et al., 1986). There are 4 *met* boxes in the *metR* promoter of *E.coli*. The sequences AGgatTtT AGcCGTCc AGAtGTtT AcACaTCc correspond to a 50%, 75%, 75%, and 63% identity to the consensus sequence.

The *met* repressor is the product of the *metJ* gene. It is a stable homodimer in dilute solution. The free repressor has a relatively low affinity for DNA. When it non-cooperatively binds two molecules of SAM, with a K_d of 10^{-5} M, it forms an active repressor that has high affinity for DNA (Old et al., 1991). Met repressor makes direct contact with the major groove in the middle of the *met* box (Phillips and Stockley, 1996).

This network is engineered from disparate elements that do not naturally interact. By introducing MetJ repression, through DNA binding, this network represents the first type of macromolecular interaction, protein:DNA. As T7-RNAP is increased, both *metJ* and GFP are transcribed (Fig. 6b). As the level of MetJ increases, SAM is cooperatively bound, increasing the affinity for the promoter's met boxes. The binding of the promoter's met boxes creates a competitive inhibition preventing T7-RNAP from transcribing the *metR* promoter-GFP hybrid. Since *metJ* and GFP do not share promoters

the reduction in read through for GFP should not reduce the quantity of MetJ transcribed. This regulated network shows a dose dependent increase/decrease in fluorescence. In Fig. 6b the relative fluorescence appears to level off and remain constant as the amount of IPTG is increased above 100 mM. We attribute this to saturation of the repression kinetics.

When we mutate the center two met boxes to the consensus sequence, the resulting region corresponds to 50%, 100%, 100%, and 63% identity. Mutation of the sequence was confirmed to increase overall repression of GFP production. In our network, increasing the repression, Fig. 6b, results in a broadening of the peak and an apparent decrease in the peak concentration without a change in the inhibitory slope. The change we see in the response as a result of these mutations confirms the model predictions.

4.3 gfp:anti-gfp network

The next network we looked at (Fig. 7a) is based on the ability of *E.coli* to utilize regulatory RNAs, also termed non-coding RNAs or small RNAs. Native regulatory RNAs have a wide variety of biological functions including the repression and activation of translation and the protection and degradation of mRNAs via base pairing with the target transcripts (Gottesman, 2004; Storz, et al., 2004). Another group of small RNAs modifies protein activity by mimicking the structures of other nucleic acids (Storz, et al., 2005).

Algorithms have predicted promoter sequences at high frequency in the *E.coli* genome, (Huerta and Collado-Vides, 2003). A whole genome expression study using oligonucleotide microarrays reported that transcripts could be detected from the antisense strand between 3000 and 4000 predicted ORFs, Selinger et al. (2000). Experimental work by Kawano et al., (2005) only obtained a limited number of clones corresponding to *cis*-encoded antisense RNAs. They suggest that most of the antisense transcripts do not persist and thus may not be significant for the cell.

The *anti-gfp* network (Fig. 7a) utilizes the production of the reverse complement of the coding GFP mRNA as an inhibitor of GFP translation. A T7 promoter downstream of the GFP coding region is oriented in such a way that the mRNA produced is the reverse compliment of GFP including the -10 region (*anti-gfp*). This network is an RNA:RNA interaction network. As the concentration of T7-RNAP increases with increasing IPTG, both GFP RNA and its reverse compliment are produced.

We have looked at the ability of this network to be tuned by changing the length of the *anti-gfp* that is transcribed. By removing half of the *anti-gfp* gene, the region that codes for the carboxy termini, we are able to see repression of the system at higher levels of IPTG that is unable to completely shut down translation of measurable GFP. In this network it appears that the overall length of the reverse complement has a dramatic effect on the inhibitory potential.

Unlike the previous network, there is no interaction between the products of T7-RNAP and T7-RNAP itself. Therefore, the kinetics of transcription should not be driving the behavior of this system (Fig. 7b). T7-RNAP efficacy is dependent to a large extent on the sequence of the binding site, which is the consensus for the forward and reverse genes; but also to a smaller extent, the sequence flanking the binding site. Differences in the flanking regions of the T7 binding site may contribute to a site strength variation that allows the system to produce an initial increase in GFP that is then returned to basal levels as both the forward and reverse T7 T7-RNAP binding sites become saturated. The signal continues to decrease at IPTG concentrations above 100 mM.

4.4 T7 T7-RNAP:T7 lysozyme Network

To improve the ability of the system to shutdown GFP production we added an additional inhibitory term to the network topology (Fig. 1b). The final network we looked at (Fig. 8a) is based on the well characterized T7-RNAP:T7 lysozyme interaction. Unlike transcriptional repressors which sterically block promoters, T7 lysozyme does not bind DNA. Rather it directly binds to T7-RNAP to form a tertiary complex with the polymerase, (Kumar and Patel, 1997), thereby modulating T7-RNAP transcriptional efficacy (Moffatt and Studier, 1987). T7 lysozyme binds a remote site and acts allosterically to effect the interactions of T7-RNAP with the non-template or the template strand, (Stano and Patel, 2004).

T7 lysozyme interacts with parts of the palm, finger, and the N-terminal domain of T7-RNAP. Protein flexibility is decreased, inhibiting a conformational change that is

required to form a fully open initiation complex (Stano and Patel, 2004). This effect can be measured as an apparent T7 lysozyme-induced increase in the K_{NTP} (Huang et al. 1999; Villemain and Sousa, 1998). Once the polymerase has cleared the promoter, the elongation complex (EC) is generally resistant to T7 lysozyme (Zhang and Studier, 1997; Kumar and Patel, 1997). However, if the RNA:T7-RNAP interaction is disrupted the EC becomes sensitive to T7 lysozyme (Huang et al., 1999)

As the concentration of T7-RNAP increases both T7 lysozyme and GFP are produced. The inhibitory effect of T7 lysozyme decreases both GFP production and T7 lysozyme production. This protein:protein interaction network produces the expected result of basal level GFP production (Fig. 8b) that responds in a dose dependent manner. The additional inhibitory connection (Fig. 1b) and improved protein stability of complexed proteins may explain the improved ability of the system to shutdown GFP production.

Two types of changes to the network were introduced to alter the characteristics and provide "tuning." First, the amount of T7 lysozyme available to the system was increased by using a plasmid with a higher copy number. By increasing the number of genes available we have been able to measure an increased output. The increase in overall repression, at a given IPTG concentration, results in a shift of the modified system's apparent peak to lower concentrations, Fig. 8b. This result is in agreement with the expected behavior of the FFL architecture seen in the previous two networks.

Next, in order to lower the amount of T7 lysozyme, the -10 Ribosome binding site (RBS) controlling T7 lysozyme was altered (Table 1). Three alternate RBSs were introduced in place of the native sequence to produce a range of -10 region strengths. Translation of T7 lysozyme is dependent on the mutated RBSs which were expected to produce two strong, one medium, and one weak variants. The functionality of each RBS was determined by measuring the fluorescent output of the RBS-GFP construct. We define translational efficiency (TE) as the ratio of the output for each RBS versus the highest output.

The resulting networks responded to changes in available T7 lysozyme in the expected fashion, Fig. 9. All of the strong RBS sites were able to reproducibly maintain the expected curve from the native network expression, with only minor variations in peak concentration and breadth. As the translation efficacy is reduced, the ability of the network to suppress fluorescence and produce an IPTG dependent decrease is diminished. Under these conditions, the data shows the medium strength RBS has a reduced ability to down regulate the system. This trend is followed with the weak RBS mutant. It appears that there is not sufficient T7 lysozyme in the system to either dampen the fluorescence or down regulate the system even at the higher levels noted.

Regulation of T7 lysozyme concentration results in the ability to alter network characteristics. Whereas, at higher concentrations, an increase in the inhibition shifts the peak to lower concentrations by altering the both the leading and lagging slope of the curve, this is not the case for low amounts of inhibitor. The amount of inhibitor (Fig. 9)

shifts the lagging slope without having a significant impact on the leading slope resulting in an overall broadening of the peak and a slight rightward shift.

5. Discussion

Basu and colleagues (2005) have reported the construction of a synthetic network that acts as a biological concentration sensor. They showed that as the concentration of inducer increases GFP expression switches from OFF to ON to OFF again. Basu's network contained five genes. In the work presented here we report three networks using a much smaller design comprised of only three genes.

The abundance of the FFL architecture in different organisms suggests that it provides a fundamental motif that can be used for more complex networks. Theoretical protein expression profiles of coherent (type 1) and incoherent (type 3, 4) FFL were computationally investigated by Ishihara et al. (2005). They show that the levels of Z monotonically change in a coherent FFL, and that monotonic and bell-shaped curves, convex or concave, can be produced depending on which type of incoherent FFL is used. Additionally they produced more complex networks, *in silico*, in order to provide a theoretical underpinning to explain experimental network behavior.

In this paper we report the construction of three different FFL networks using different biological components and show that they have equivalent qualitative characteristics, though varying quantitative responses. All responses were measured at steady state as confirmed by tracking dGFP/dt (production minus degradation) over time (Fig. 4). In

addition to experimental work, we also constructed a very simple model of a generic FFL network. Using this model we could simulate the effects of different mutations on the behavior of the network. In particular, mutations at the inhibitory binding site and copy number result in changes to the peak location and height. In order to verify the model predictions we investigated the effect of mutating the inhibitory elements on the peak position for each network.

In the MetJ:metR promoter network the ability of MetJ to inhibit T7-RNAP by binding the *metR* promoter region was mutated from the native sequence to one that had two consensus boxes. This change produces a binding site that has a higher affinity for MetJ and therefore acts as a stronger repression signal. When we investigated this change, the increase in the inhibition shifted the peak concentration to the left, as predicted in our simulations.

The *gfp:anti-gfp* network was mutated to decrease the amount of inhibition present. We expected that decreasing the size of the *anti-gfp* produced would result in a decrease in inhibition. This is seen experimentally by the increase in the peak concentration and decrease in the system's ability to shut down GFP translation.

Finally, the concentration of T7 lysozyme was modulated by increasing and decreasing the amount of translated inhibitor. In the former case, an increase in the copy number results in an eight fold increase in available T7-lysozyme. The increase in expression was quantitated by utilizing the ratio of the normalized fluorescence produced by GFP

under the same conditions. The increased repression by T7-lysozyme produces a leftward shift of the peak. In the latter case, by reducing the ability of the transcribed T7 lysozyme to produce inhibitor by mutating the ribosome binding sequence, we were able to down-regulate the inhibitory element. This produced a broadening and a slight rightward of the peak. In each case we saw the expected change in the peak position, supporting the predictions made by the model. In principle this would allow an engineer to tune the network to a desired behavior.

Ultimately one would like to combine sub-networks such as those described here into larger networks to generate more complex behaviors such as timed multiple events, oscillators or homeostatic devices. The FFL is sufficiently flexible to enable such networks to be generated (unpublished). However in order to connect different FFL sub-networks together would require a redesign of the current set. One of the main problems is that the networks reported here share common components, for example the T7 promoter is used in all the networks. Connecting multiple FFL motifs together would result in cross-talk between the sub-networks. To avoid this we need to utilize unique components for each FFL network. A second problem is that the different FFL sub-networks need to be interfaced to each other, currently the FFL networks reported here use a common input component, this would have to be changed in each network to enable connections to be made, this applies equally to the GFP output protein. We believe such alterations could be made to the sub-networks so that larger networks can be constructed. What is most promising is that we found the networks to be fairly robust and operated

immediately as expected when construction, we did not, for example, need to tune or adjust conditions in order for the networks to operate.

Future work will entail the construction of plug-and-play FFLs. Unpublished theoretical work that we have done has shown that FFLs are behaviorally very flexible units and can be used to build a variety of interesting devices. Once we have constructed plug-and-play units we intend to build more complex devices.

Acknowledgements

This work was supported by a grant from the National Science Foundation (Id 0432190). The GFP was kindly donated by Drew Endy form the Biobricks repository. We would also like to acknowledge useful discussions with Vijay Chickarmane.

References

Ackers, G.K., Johnson, A.D., and Shea, M.A. (1982). Quantitative Model for Gene-Regulation by Lambda-Phage Repressor. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **79:** 1129-1133.

Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H., and Weiss, R. (2005). A synthetic multicellular system for programmed pattern formation. *Nature* **434:** 1130-1134.

Basu, S., Mehreja, R., Thiberge, S., Chen, M.T., and Weiss, R. (2004). Spatiotemporal control of gene expression with pulse-generating networks. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 6355-6360.

Belfaiza, J., Parsot, C., Martel, A., Delatour, C.B., Margarita, D., Cohen, G.N., and Saintgirons, I. (1986). Evolution in Biosynthetic Pathways - 2 Enzymes Catalyzing Consecutive Steps in Methionine Biosynthesis Originate from a Common Ancestor and Possess a Similar Regulatory Region. *Proceedings of the National Academy of Sciences of the United States of America* **83:** 867-871.

Bray, D. 1995. Protein Molecules as Computational Elements in Living Cells. *Nature* **376:** 307-312.

Gottesman, S. (2004). The small RNA regulators of Escherichia coli: Roles and mechanisms. *Annual Review of Microbiology* **58:** 303-328.

Hartwell, L.H., Hopfield, J.J., Leibler, S., and Murray, A.W. (1999). From molecular to modular cell biology. *Nature* **402:** C47-C52.

Huang, J.B., Villemain, J., Padilla, R., and Sousa, R. (1999). Mechanisms by which T7 lysozyme specifically regulates T7 RNA polymerase during different phases of transcription. *Journal of Molecular Biology* **293:** 457-475.

Huerta, A.M., and Collado-Vides, J. (2003). Sigma70 promoters in Escherichia coli: Specific transcription in dense regions of overlapping promoter-like signals. *Journal of Molecular Biology* **333:** 261-278.

Ishihara, S., Fujimoto, K., and Shibata, T. (2005). Cross talking of network motifs in gene regulation that generates temporal pulses and spatial stripes. *Genes to Cells* **10:** 1025-1038.

Kaern, M. and Weiss, R., Synthetic gene regulatory systems. In Jrg Stelling Zoltan Szallasi and Vipul Periwal, editors, *System Modeling in Cellular Biology From Concepts to Nuts and Bolts*, (2006), MIT Press, chapter 13, pages 269{298.

Kawano, M., Reynolds, A.A., Miranda-Rios, J., and Storz G. (2005) Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in Escherichia coli, Nucleic Acids Res **33**: 040-1050.

Kumar, A., and Patel, S.S. (1997). Inhibition of T7 RNA polymerase: Transcription initiation and transition from initiation to elongation are inhibited by T7 lysozyme via a ternary complex with RNA polymerase and promoter DNA. *Biochemistry* **36:** 13954-13962.

Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., et al. (2002). Transcriptional regulatory networks in Saccharomyces cerevisiae. *Science* **298:** 799-804.

Mangan, S., and Alon, U. (2003). Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 11980-11985.

Mayya, V., and Loew, L.M. (2005). STAT module can function as a biphasic amplitude filter. *Systems Biology* **2:** 43-52.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science* **298:** 824-827.

Moffatt, B.A., and Studier, F.W. (1987). T7 Lysozyme Inhibits Transcription by T7 Rna-Polymerase. *Cell* **49:** 221-227.

Old, I.G., Phillips, S.E.V., Stockley, P.G., and Saintgirons, I. (1991). Regulation of Methionine Biosynthesis in the Enterobacteriaceae. *Progress in Biophysics & Molecular Biology* **56:** 145-185.

Phillips, S.E.V., and Stockley, P.G. (1996). Structure and function of Escherichia coli met repressor: Similarities and contrasts with trp repressor. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **351:** 527-535.

Sauro, H.M. and Kholodenko B.N. (2004) Quantitative analysis of signaling networks, Prog Biophys Mol Bio **86**: 5-43.

Sauro H.M, Hucka M, Finney A, Wellock C, Bolouri H, Doyle J and Kitano H. (2003) Next generation simulation tools: the Systems Biology Workbench and BioSPICE integration. OMICS. 2003 Winter;7(4):355-72.

Selinger, D.W., Cheung, K.J., Mei, R., Johansson, E.M., Richmond, C.S., Blattner, F.R., Lockhart, D.J., and Church, G.M. (2000). RNA expression analysis using a 30

base pair resolution Escherichia coli genome array. *Nature Biotechnology* **18:** 1262-1268.

Setty, Y., Mayo, A.E., Surette, M.G., and Alon U. (2003) Detailed map of a cis-regulatory input function, P Natl Acad Sci USA **100**: 7702-7707.

Shen-Orr, S.S., Milo, R., Mangan, S., and Alon, U. (2002). Network motifs in the transcriptional regulation network of Escherichia coli. *Nature Genetics* **31:** 64-68.

Stano, N.M., and Patel, S.S. (2004). T7 lysozyme represses T7 RNA polymerase transcription by destabilizing the open complex during initiation. *Journal of Biological Chemistry* **279:** 16136-16143.

Storz, G., Altuvia, S., and Wassarman, K.M. (2005). An abundance of RNA regulators. *Annual Review of Biochemistry* **74:** 199-217.

Storz, G., J.A. Opdyke, and Zhang A.X.(2004) Controlling mRNA stability and translation with small, noncoding RNAs. Current Opinion in Microbiology **7**: 140-144.

Storz, G., Opdyke, J.A., and Zhang, A.X. (2004). Controlling mRNA stability and translation with small, non-coding RNAs. *Current Opinion in Microbiology* **7:** 140-144.

Villemain, J., and Sousa, R. (1998). Specificity in transcriptional regulation in the absence of specific DNA binding sites: The case of T7 lysozyme. *Journal of Molecular Biology* **281:** 793-802.

Weissbach, H., and Brot. N. (1991) MicroReview Regulation of methionine synthesis in *Escherichia coli. Molecular Microbiology* **5(7),** 1693-1597.

Wolf, D.M., and Arkin, A.P. (2003). Motifs, modules and games in bacteria. *Current Opinion in Microbiology* **6:** 125-134.

Zhang, X., and Studier, F.W. (1997). Mechanism of inhibition of bacteriophage T7 RNA polymerase by T7 lysozyme. *Journal of Molecular Biology* **269:** 10-27.

Figure Captions

Fig. 1 | The Incoherent Feed-forward Loop Type 1.

The Incoherent Feed-forward Loop Type 1 was used as the basis for designing networks. Two Networks were designed that fit this scheme, where X is T7 RNA polymerase and Z is GFP. (A) Y is the inhibitory element, MetJ or the reverse complement of GFP. (B) This network is a modified FFL with additional feedback, Y is T7 lysozyme.

Fig. 2 | Feed-forward Loop Model.

Interpretation of the feed-forward motif shown (Fig. 1a) in terms of a very simple genetic network. p1 is the input to the system comprising of genes G1 and G2, which produce p2 and p3 respectively. p1 activates G2 and p2 represses G2.

Fig. 3 | Model Predictions.

Steady state values of p3 for the simple generic incoherent type I feed-forward network as a function of the input p1 for the model shown in the Fig. 2. The bold line represents the 'wildtype' model. The dotted line represents an increase in the affinity of p2 for G2. The dashed line represents a further increase in the affinity of p2 for G2. Parameter values for the equations shown in the modeling section are set to: t1 = 5; a1 = 0.1; t2 = 1; b1 = 1; b2 = 0.1. The curves were generated by changing the affinity parameter b3 on G2 from 1 to 5 to 20, where 1 is the wild-type value. Gamma_1 and Gamma_2 were set to 0.1. Note that the curves have all been normalized with respect to the maximum peak high for each curve. This was done to emphasize the lateral change in the response curve

which was measurable in our experiments. For the unnormalized curves, the peak height decreases as the inhibition strength increases as well as shifting the curve to the left.

Fig. 4 | Steady State.

This plot shows the rate of change of GFP concentration (production minus degradation) normalized over OD600 as a function of time for four representative runs. Each network tested in this paper was tracked over time to verify that the networks had reached their steady state operating points. A rate of change of near zero indicates that the synthetic networks have reached steady state (approximately after 16hrs to 20 hrs).

Fig. 5 | Unregulated System.

The partial networks, without inhibition, used in this work were independently characterized for IPTG dose dependent responses. (A) GFP production is driven by T7-RNAP binding approximately 70 nucleotides upstream of the GFP start codon. (B) GFP production with a met R promoter insert. A T7-RNAP binding site was introduced directly upstream of the complete met R promoter region. (C) A plot of the normalized fluorescence vs. IPTG concentration. Fluorescence was measured for each of the constructs, normalized by the optical density of each well, and scaled according to each construct's maximal output. Increasing IPTG induces the production of T7-RNAP which in turn transcribes the downstream GFP gene: T7 promoter (-●-); T7-*metR* promoter (-■-). Curves were fitted to a simple Hill function using EasyGraph (Future Skill Software) and the Levenberg-Marquardt optimization algorithm. Hill coefficients for the T7 promoter (-●-) equaled 1.93 while for the *metR* promoter (-■-) it equaled 1.42

Fig. 6 | T7-*metR* promoter Regulated System.

The complete network with MetJ inhibition (Degradation steps have been omitted for clarity). (A) GFP and MetJ production are independently driven by T7-RNAP binding. MetJ inhibits by binding "met boxes" located in the regulatory region of the *metR* promoter driving GFP production. (B) A plot of the normalized fluorescence vs. IPTG concentration. Fluorescence was measured for each of the constructs and normalized by the optical density of each well, (-●-) native *metR* promoter region, (-■-) *metR* promoter where the center two met boxes have been mutated to the consensus sequence.

Fig. 7 | T7-*anti-gfp* Regulated System

The complete network with *anti-gfp* inhibition (Degradation steps have been omitted for clarity). (A) The reverse complement of the GFP is transcribed by a T7 binding. The *anti-gfp* transcript encodes the entire gene and upstream ribosome binding site. Inhibition is through mRNA:mRNA binding of the GFP gene and its reverse complement to prevent translation. (B) A plot of the normalized fluorescence vs. IPTG concentration. The entire reverse complement of the (-■-) GFP gene and a (-●-) partial sequence were measured for the ability to inhibit.

Fig. 8 | T7-T7 lysozyme Regulated System

Simplified representation of the T7 Lysozyme inhibition System (Degradation steps have been omitted for clarity). (A) GFP and T7 lysozyme production are independently driven by T7-RNAP binding. T7 lysozyme inhibits by binding T7-RNAP and preventing

elongation complex formation. (B) A plot of the normalized fluorescence vs. IPTG concentration. The introduction of (-●-) lysozyme repression and (-■-) an eight fold increase T7 lysozyme, available to the system through gene duplication, over the native system, was measured.

Fig. 9 | Tuning T7-T7 lysozyme Regulated System

The lysozyme inhibitory network was mutated to decrease the amount of lysozyme present without changing T7-RNAP kinetics. The ribosome binding site (RBS) was altered, from strong to weak, to decrease mRNA translation into active inhibitor. The ability of the T7 lysozyme to effectively inhibit GFP production is proportional to RBS Strength. The RBS strengths are as follows: strongest (-♦-), strong (-■-), medium (-●-), weak (-▲-)

Table 1 | Ribosome Binding Sequence.

The ribosome binding sequence controlling the production of T7-lysozyme was mutated to decrease the overall efficiency of translation. In order to quantitate the sequences, the output of each RBS driving the production of GFP was determined. We define translational efficiency (TE) as the ratio of the output for each RBS versus the highest output. The nucleotides in bold indicate the RBS and the underlined sequence is the protein start sequence.

A.

B.

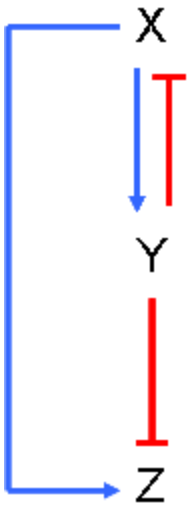


Fig. 1

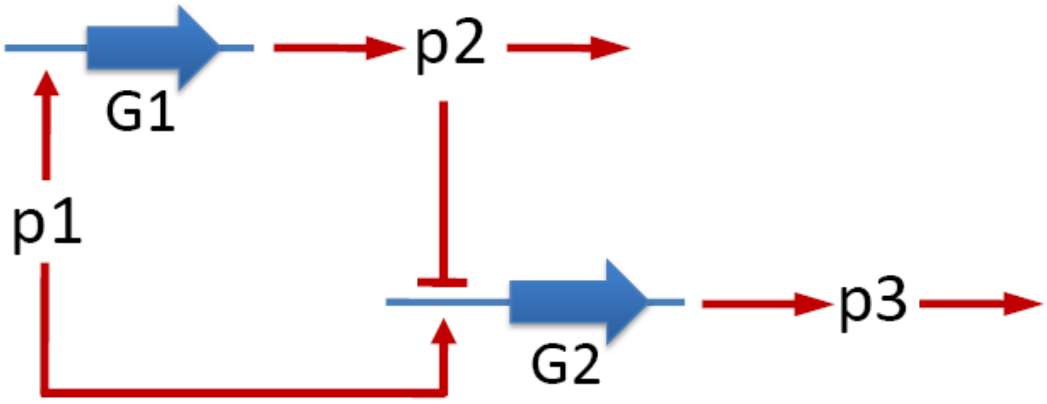


Fig. 2

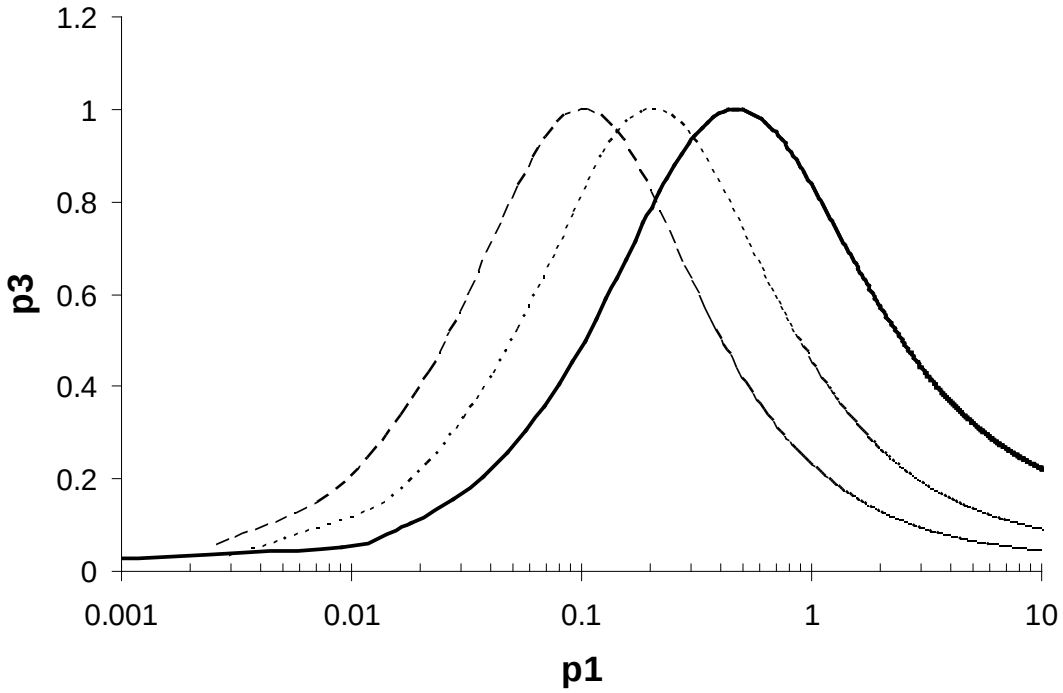


Fig. 3

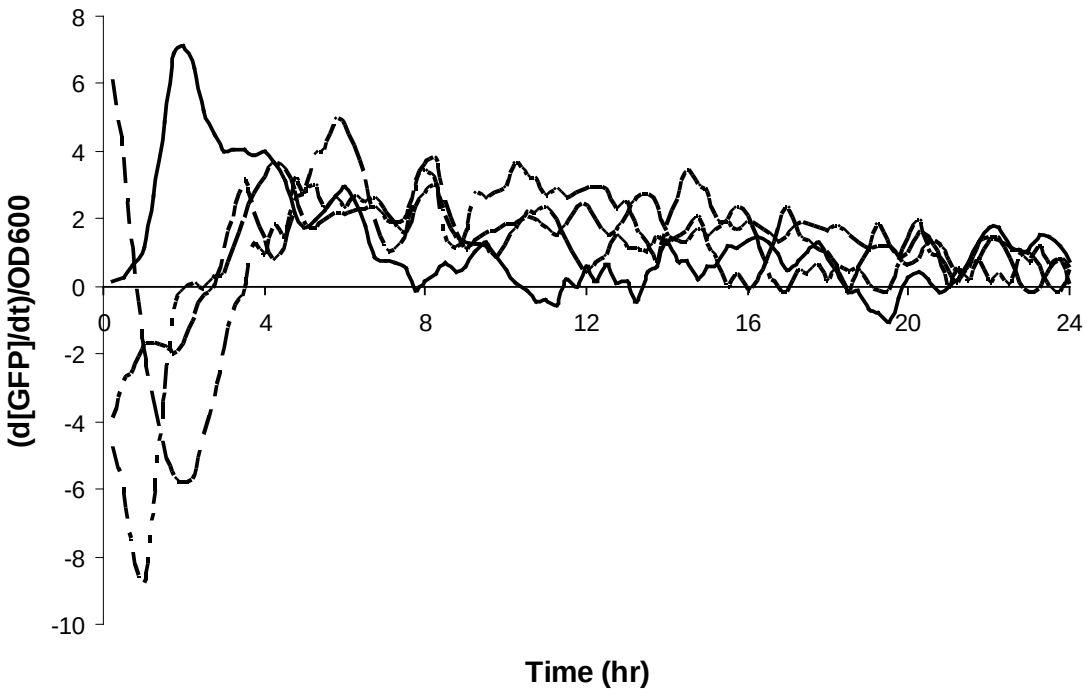
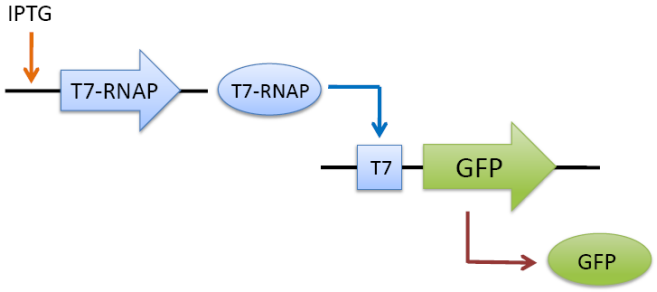
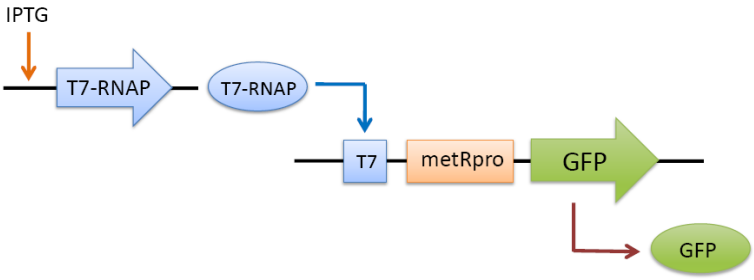


Fig. 4

A.


IPTG
T7-RNAP
T7-RNAP
T7
GFP
GFP

B.


IPTG
T7-RNAP
T7-RNAP
T7
metRpro
GFP
GFP

C.

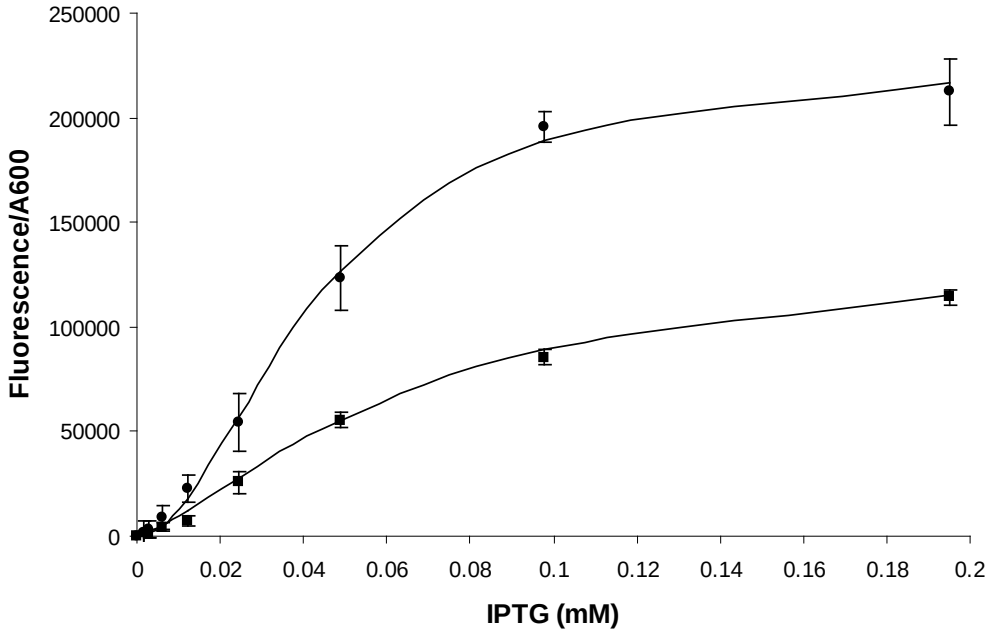
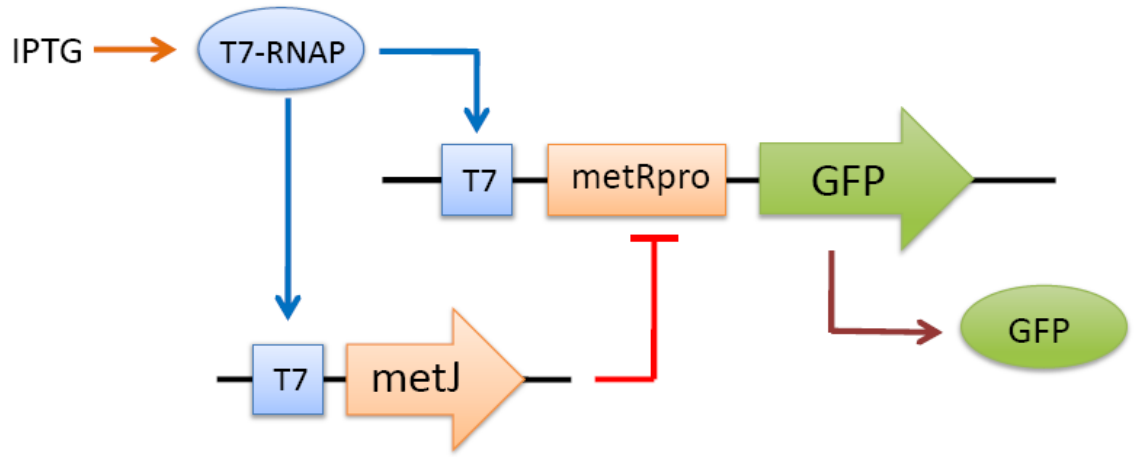

250000
200000
150000
100000
50000
0
Fluorescence/A600
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.2
IPTG (mM)

Fig.5

A.


IPTG
T7-RNAP
T7
metRpro
GFP
GFP
T7
metJ

B.

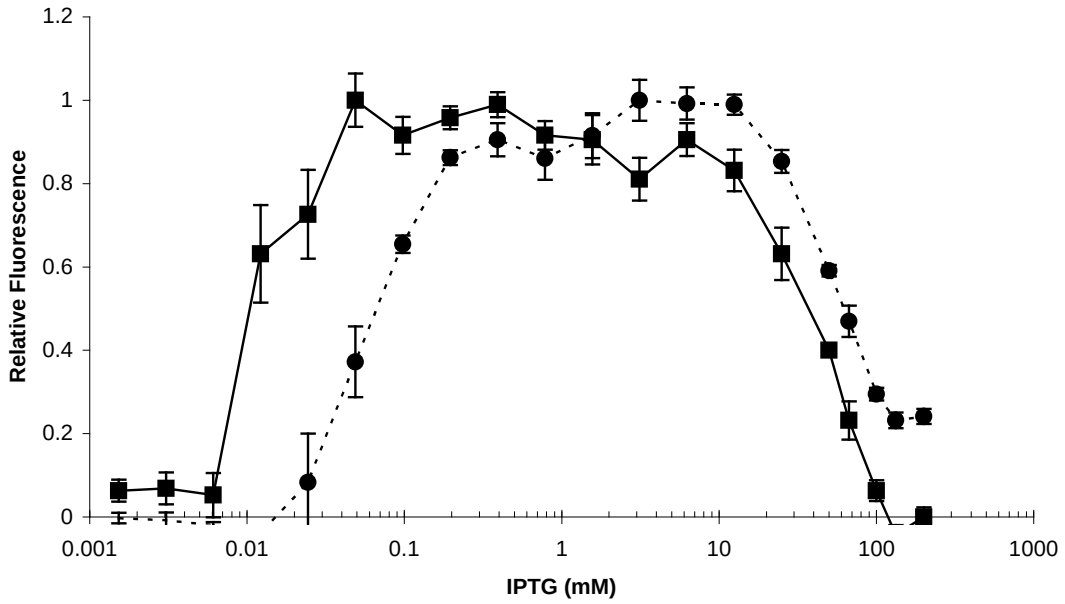
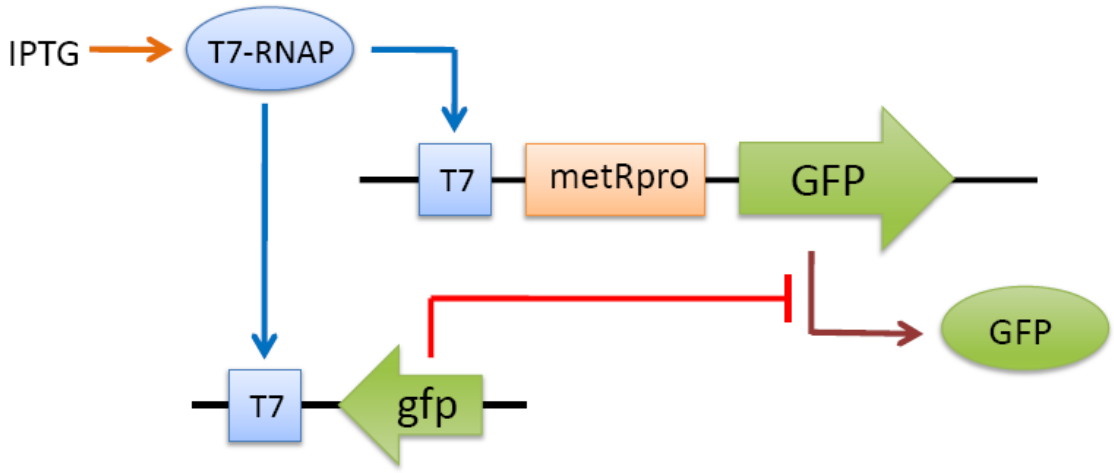

Relative Fluorescence
1.2
1
0.8
0.6
0.4
0.2
0
0.001
0.01
0.1
1
10
100
1000
IPTG (mM)

Fig. 6

A.


IPTG
T7-RNAP
T7
metRpro
GFP
GFP
T7
gfp

B.

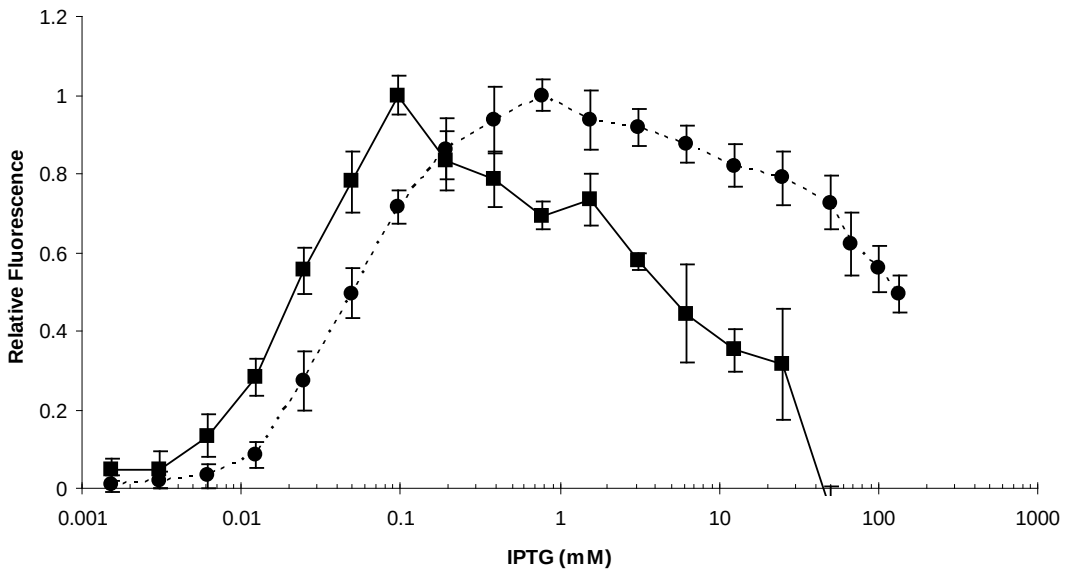
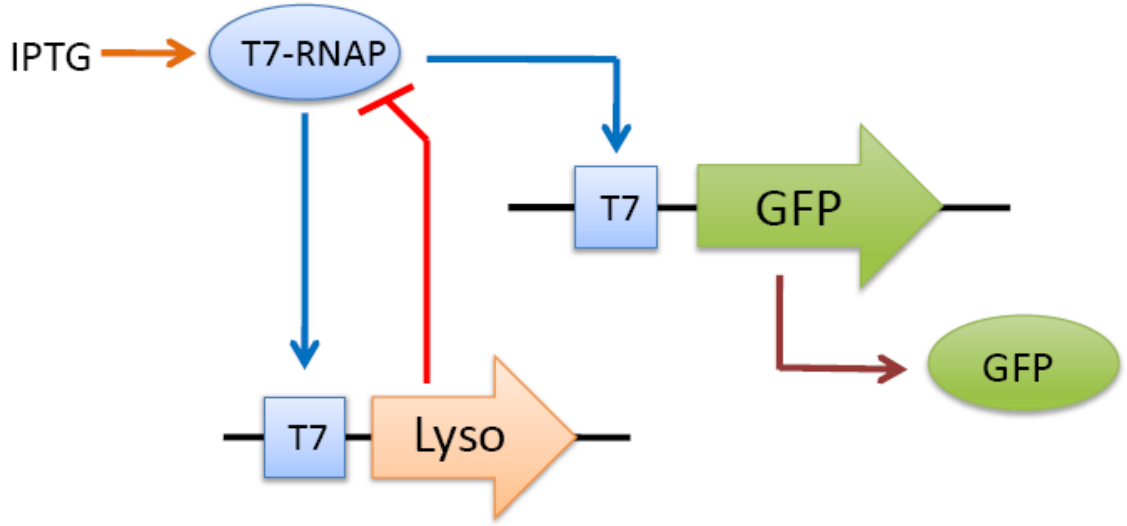

Relative Fluorescence
1.2
1
0.8
0.6
0.4
0.2
0
0.001
0.01
0.1
1
10
100
1000
IPTG (mM)

Fig. 7

A.


IPTG
T7-RNAP
T7
GFP
GFP
T7
Lyso

B.

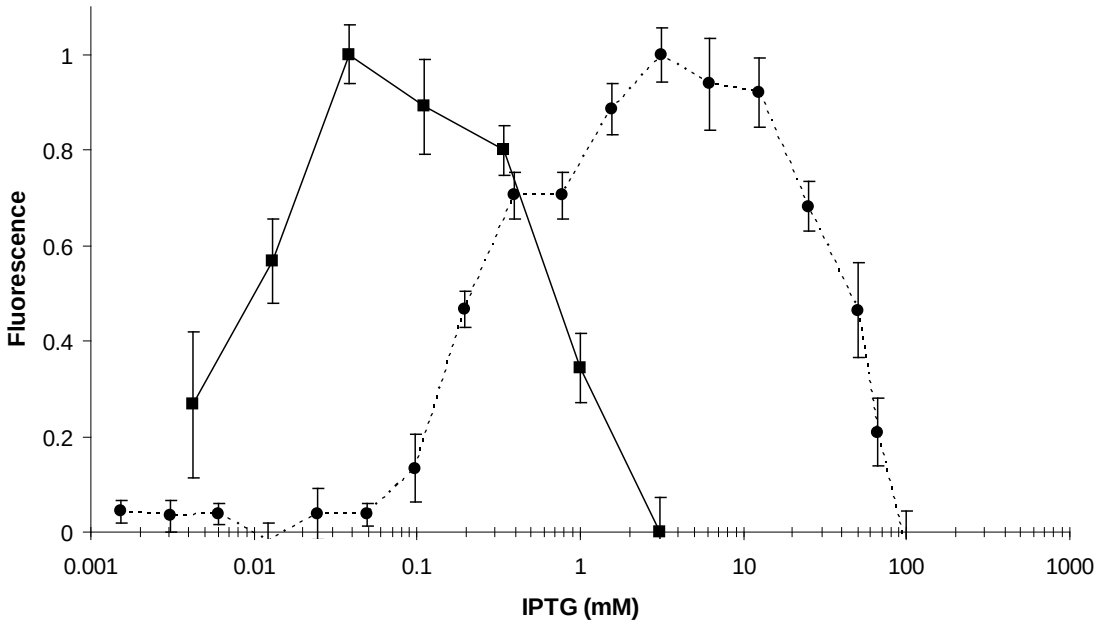

Fluorescence
1
0.8
0.6
0.4
0.2
0
0.001
0.01
0.1
1
10
100
1000
IPTG (mM)

Fig. 8

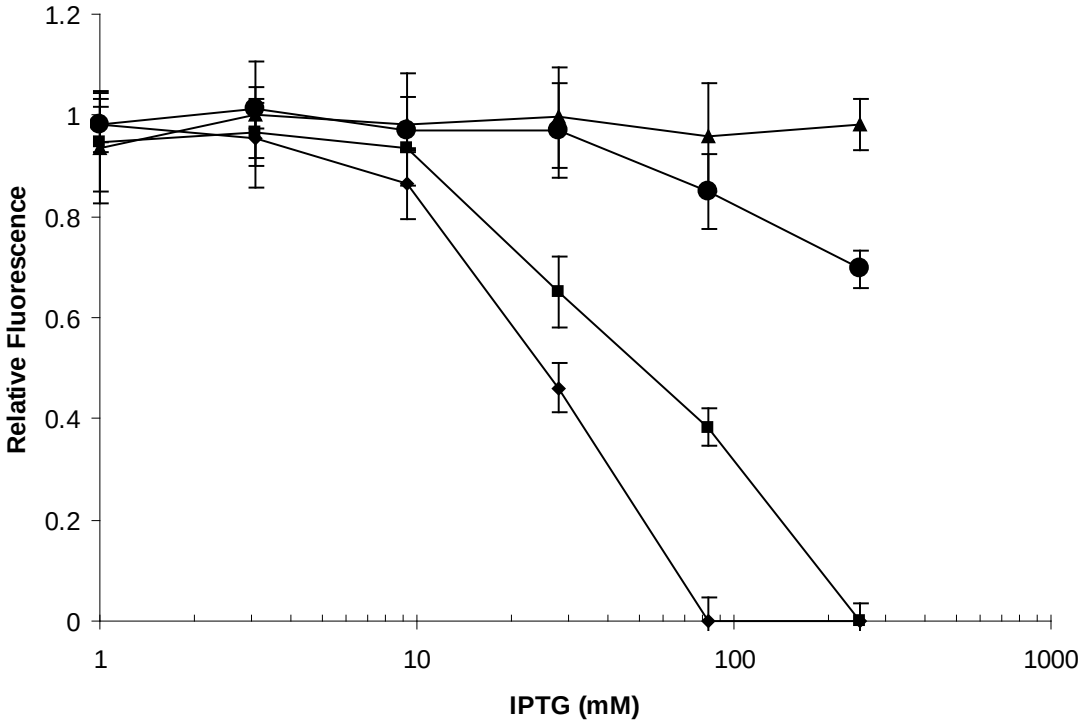


Fig. 9

	RBS Sequence	TE
1.	AAG**AAGGAG**ATATACCATG	1.00
2.	TAAG**AAGGAA**ATTAATCATG	0.95
3.	AACA**CAGGAA**AATTAATCATG	0.60
4.	AACAC**AGGAAC**AATTAATCATG	0.45