# Bi 1x
# A Synthetic Delay Oscillator

## 1 Objective

In this lab, you will analyze a synthetically engineered bacterial oscillator, one of many circuits that have been implemented *in vivo*. Using skills developed from prior modules featuring high-magnification phase and fluorescent microscopy with time-lapse imaging, you will observe and quantify the dynamics of a single-gene delay oscillator in *E. coli*. With image segmentation and tracking tools, coupled with mathematical modeling, you will be able to directly compare a theoretical description of protein production, degradation and repression with experimental measurements.

## 2 Background

### 2.1 Modular components in synthetic biology

Synthetic biology can be generally summarized as: (1) using existing components of biology as a building block, or (2) using biology as a motivation to engineer new systems, together improving nature's networks and their functions. Synthetic biology requires design and modeling to develop predictably functional and complex systems from smaller parts and devices known as core components. Integrated circuit designs in electrical engineering have a vastly large and defined set of components, whose material properties are well characterized, allowing for extremely predictable and reliable functioning circuits (Figure 1a). The final processing power of integrated circuits correlate directly to the modeled features, allowing for tunable performance. With recent advances in biology, we have identified and developed a library of standardized parts and their performance (Figure 1b). Such efforts promote the incorporation of traditional engineering design principles, advancing robust circuit design in synthetic biology.

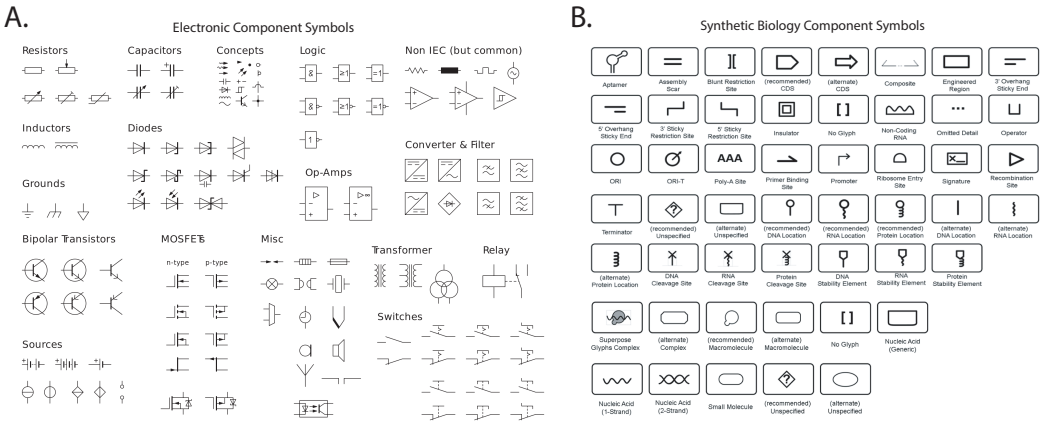


Figure 1: **a.** Standard graphic symbols of individual components used in electronic and electrical engineering diagrams. **b.** Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology.

Similar to electronic circuit designs, synthetic biology has developed a standard for design components known as *"Synthetic Biology Open Language"* (SBOL). The SBOL is composed of a library of reusable components that can be hierarchically placed together to build synthetic biological circuits. For an effective library system of robust parts, the supplier of DNA components libraries, the designer who uses the components in the library must speak the same language and use the same terminology.

Plasmids are commonly used in recombinant DNA technology. In *E. coli*, plasmids are relatively small, approximately 3000–6000 basepairs (and with recent advancements in DNA assembly, larger plasmids are now possible), typically containing only the essential DNA sequences for replication; a DNA origin of replication, an antibiotic-resistance gene, and a region in which exogenous DNA fragments can be inserted. Each part can be assembled through Golden-Gate or Gibson assembly which enables quick and robust customizable function. Extrachromosomal plasmids in *E. coli* replicate during cell division and are segregated to the resulting daughter cells.

When engineering synthetic biological systems for *in vivo* control, it is important to first determine the host organism of interest and identify the available modular components. Designing synthetic *in vivo* systems with predictable functionality requires modular and orthogonal plasmid elements that are tunable at multiple layers. The first level of control is the **origin of replication (ORI)**. The ORI is a DNA sequence that directs initiation of plasmid replication by recruiting the necessary transcriptional machinery. The properties of the ORI determine how many plasmid copies are in each cell. Changing the number of copies within a cell can affect the

transcriptional efficiency as well as the target gene expression rates. The trancriptional unit is composed of multiple tunable components, housing the target gene, starting with the promoter domain and ending with the terminator (Fig. 2).

The first regulatory element within the transcriptional unit is the **promoter**. The promoter sequence contains a binding domain that controls the binding of the RNA polymerase and other transcriptional factors necessary to begin transcription. The binding affinity of various promoters plays a large role in determining how strongly the target gene will be expressed. Promoters can also be controlled by the presence of molecular biology reagents; promoter repression occurs by a chemical compound sequestering binding domain preventing binding of polymerases, and activation usually via a secondary inducer that displaces original repressing chemical. These promoters are called inducible promoters. In this module, we will consider dynamics of **repressors**, treRL and lacI, which bind the promoter to shut down gene expression.

Translational regulation of gene expression can be controlled by varying the **Ribosome Binding Site** (**RBS**). After transcription, where the DNA is converted to mRNA, translation initiation in bacteria requires an RBS sequence followed by a start codon (**AUG**). Similar to the promoter, the RBS has a defined domain that allows for binding of the ribosome and recruitment of all other machinery. The RBS is typically found after the promoter and before the protein coding sequence. Ribosome binding rate can be increased by minimizing the formation of secondary structures near the RBS by adding more adenines upstream of the RBS.

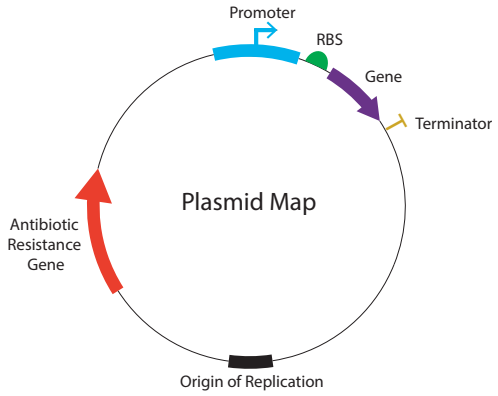


Figure 2: Sample plasmid map with core components for modular functional design. The transcriptional unit contains a promoter, RBS, gene, and terminator. The antibiotic resistance gene is present to allow screens for only cells that have the plasmid by exposing the cells to antibiotic.

When designing a plasmid, there three main levels of control; (1) Origin of replication, (2) Promoter, and (3) Ribosome binding site. In order to determine the opti-

mal combination of these three, understanding the factors such as potential protein toxicity, *E. coli* strain, and growth conditions, could dramatically affect the overall function.

Questions to consider when engineering a biological system include what will be the chassis of choice for engineering (bacteria, fungi, mammalian, DNA, etc)? What are the genetic parts parts available for the chosen chassis for circuit design? How do we integrate such building blocks of parts so that their performance can be genetically predicable?

## 2.2 Oscillators in synthetic biology

In general, oscillators are great systems to explore synthetic biological question through the design and implementation of control function.

### 2.2.1 The repressilator

One of the earliest functioning synthetic genetic circuits is an oscillator known as the **repressilator**, developed by Michael Elowitz and Stan Leibler. The repressilator consists of three repressors on a plasmid, as shown in Fig. 3. They are TetR, λ cI, and LacI.

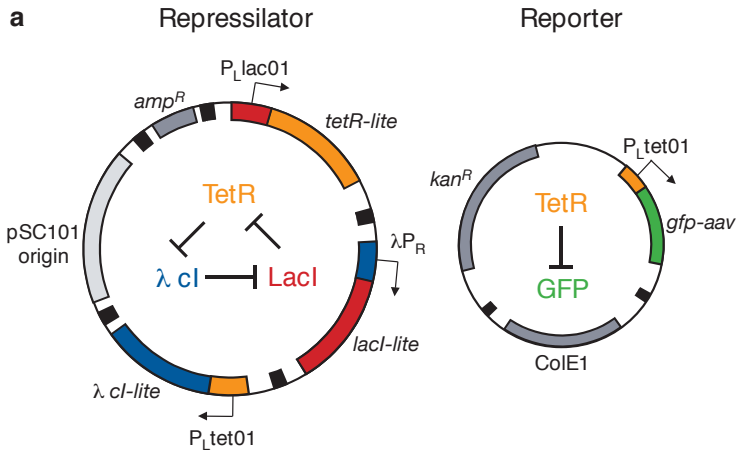


Figure 3: Schmematic of the plasmids used to construct the repessilator in *E. coli*. The three repressors TetR, λ cI, and LacI, are on the same plasmid and form a cycle of repression. Additionally, TetR represses GFP, which is found on another plasmid. The *lite* suffix on the repressors signifies that they have a destruction tag to decrease their stability. The *aav* suffix on the GFP indicates that it is a variant of intermediate stability. Taken from Elowitz and Leibler, *Science*, **403**, 335–338, 2000.

We might work out the dynamics of this system by reasoning. We might get a stable steady state, where all three levels of repressor are tuned to reasonably repress

the others. Conversely, we might image a dynamic scenario. Say that initially TetR has high copy number and $\lambda$ cI and LacI are low. The high copy of number of TetR will keep the numbers of $\lambda$ cI down. This means that LacI is free to be expressed. As its copy number grows, it will start to repress TetR. As TetR goes down, $\lambda$ cI is expressed in higher numbers. The increased $\lambda$ cI copy number leads to less LacI. Then, TetR comes back up again. So, we see a cycle, where TetR is high, then LacI, and finally $\lambda$ cI. Since TetR represses the fluorescent protein GFP, we will see an oscillation in GFP as TetR goes up and down. In this case, GFP is the **reporter**.

We may also think of the repressilator as follows. TetR is **autorepressive** in the sense that it is in a **negative feedback loop**. It indirectly represses itself, by repressing the gene that represses the gene that represses itself. Its expression oscillates because there is a **delay** between when its expression level rises and when the repression takes effect. We must wait for $\lambda$ cI to be repressed and then for LacI to rise.

### 2.2.2 Delay oscillators

In thinking about the repressilator, we see that the two key features to get oscillations are negative feedback and delay. We can conceive a simple single-gene oscillator if we have an autorepressive gene. The principle behind the presence of oscillations is simple, and illustrated in Fig. 4. Because the rate of expression is determined by past protein levels, high expression occurs when protein levels were low in the past and low expression occurs when protein levels were high in the past. Synthetic delay oscillators have been constructed; see, for example, Stricker, et al., *Nature*, **456**, 516–519, 2008.

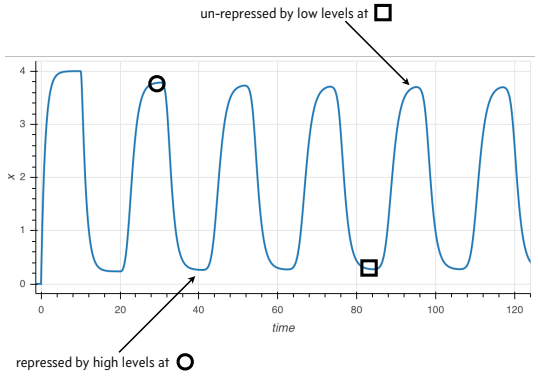


Figure 4: Schematic of how delayed repression can give oscillations.

## 2.3 The Swaminathan oscillator

Anandh Swaminathan in Richard Murray's lab here at Caltech created an as-of-yet unpublished synthetic single-gene delay oscillator, shown in Fig. 5. The *E. coli*

cells have two plasmids. One has a repressor, either treRL or lacI. The repressor is under the control of the PlacO promoter. The second has a reporter, which codes for a fluorescent protein, Superfolder Yellow Fluorescent Protein (sfYFP). The repression activity of the repressor occurs after a time delay, as the repressor only functions in as a multi-molecular complex, which takes time to form and then act on the promoter. The repressor acts on the PlacO promoter to downregulate its own expression and that of the dfYFP reporter. The cells also express clpXP, which enzymatically degrades both the repressor and sfYFP (which is adapted to have a tag to make is susceptible to ClpXP.)

As we will see in the mathematical modeling that follows and in the assignment, the active degradation of both the repressor and the reporter speeds up the dynamics of the oscillator. This is to date the fastest known genetic oscillator, and you will explore and quantify its dynamics in this lab. You will acquire fluorescence time-lapse movies of the Swaminathan oscillator with a treRL repressor and with a LacI repressor. You will quantify the period of the oscillators and compare the performance of the two repressors.

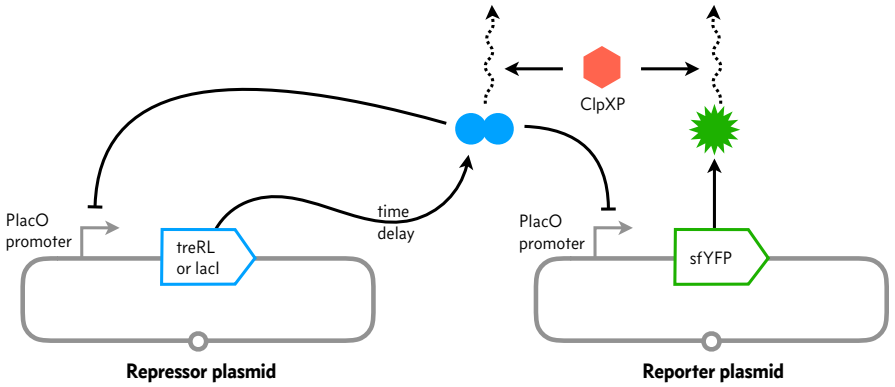


Figure 5: The delay circuit, designed and constructed by Anandh Swaminathan, that we will use in lab. Arrows indicate activation, with a dashed arrow indicating a degradation process. Blunt arrows indicate repression. Not shown are a gene imparting resistance to Kanamycin on the repressor plasmid and a gene imparting resistance to Carbenicillin on the reporter plasmid.

### 2.3.1 Modeling the oscillator

To model the Swaminathan oscillator, we will mathematize the cartoon representation in Fig. 5. We want to write differential equations describing the dynamics of the two species of interest, repressor (treRL or LacI) and sfYFP. We denote their concentrations by $x$ and $y$, respectively. So, we wish to specify the right-hand-sides of

$$\frac{dx}{dt} = \cdots , \tag{1}$$

$$\frac{dy}{dt} = \cdots . \tag{2}$$

As the cells grow, the concentrations of both the repressor and the sfYFP will drop as a result of dilution. The growth is exponential, so we model the dilution effect as

$$\frac{dx}{dt} = -\gamma_0 x + \cdots , \tag{3}$$

$$\frac{dy}{dt} = -\gamma_0 y + \cdots , \tag{4}$$

where the $\cdots$ represent terms we have not yet worked out.

In addition to dilution, the molecules are produced by gene expression. The rate of production of the repressor and the reporter per plasmid should be approximately equal, since they are under control of the same promoter. The differences in production rate are due to differing copy numbers of the plasmids. Let $R$ be the number of repressor plasmids in a cell and $F$ be the number of reporter plasmids in a cell. The respective maximum production rates are then $\beta R$ and $\beta F$. The production rate is modulated by the presence of the repressor. **Hill functions** are ofter used to describe the effect of repression. A repressive Hill function has the form

$$f(x) = \frac{1}{1 + (x/k)^n}, \tag{5}$$

where $k$ is referred to as the Hill $k$ and $n$ as the Hill coefficient. Note, though, that there is a delay in the action of the repressor. If the time delay is $\tau$, it is the concentration of repressor at time $t - \tau$ that affects expression at time $t$. With this in mind, we can add production to the dynamical equations.

$$\frac{dx}{dt} = \frac{\beta R}{1 + (x(t - \tau)/k)^n} - \gamma_0 x + \cdots , \tag{6}$$

$$\frac{dy}{dt} = \frac{\beta F}{1 + (x(t - \tau)/k)^n} - \gamma_0 y + \cdots . \tag{7}$$

Finally, we have enzymatic degradation. Many enzymatic reaction obey **Michaelis-Menten** kinetics. The expression for the enzymatic degradation can be derived from mass action kinetics, with a pseudo-steady-state assumption, considering the chemical reaction

$$\mathrm{E} + \mathrm{S} \rightleftharpoons \mathrm{ES} \rightarrow \mathrm{E}. \tag{8}$$

Here, the enzyme, E, reversibly binds its substrate (in our case, treRL, LacI, or sfYFP), and then degrades the substrate, leaving the enzyme to be recycled. Using the Michaelis-Menten expression for the degradation kinetics, we complete our

dynamical system.

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\beta R}{1 + (x(t-\tau)/k)^n} - \gamma_0 x - \gamma_1 \frac{x}{K_m + x + y}, \tag{9}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\beta F}{1 + (x(t-\tau)/k)^n} - \gamma_0 y - \gamma_1 \frac{y}{K_m + x + y}, \tag{10}$$

where we have introduced the Michaelis constant, $K_m$.

In the Jupyter notebook accompanying this module, we demonstrate how to numerically solve this system of delay differential equations (DDE). You will see in your assignment that mathematical modeling can provide insight on how to design genetic circuits and what kind of performance we can expect out of them. The notebook also contains parameters suggested by Anandh Swaminathan for the dynamics of this system.

# 3 Protocol

Prior to your involvement:

1. *E. coli* strains were inoculated overnight at 37°C in 5 mL cultures in M9CA media (Teknova M8010) in Carbenicillin [100 $\mu$g/mL] and 50% Kanamycin [25 µg/mL].
2. The following morning, outgrowths were diluted by a factor of 1000 to 5000 fold into fresh 5 mL M9CA media (containing the same concentration antibiotics as initial outgrowths). Outgrowth was performed for 4-7 hours until the cells reached an $OD_{600}$ of at least 0.1 and no more than 0.4.

## 3.1 Preparing the agar pads

The agar pads will be prepared by your TAs, but the protocol for making them is displayed below. Each group should have two or three pads, just in case there is a problem with one.

1. Stretch a piece of Parafilm on the bench. Place an 18×18 mm coverslip on the Parafilm.
2. Pipette about 500-750 µL of M9CA media with 2-3% agarose.
3. Drop a second coverslip on top in order to "sandwich" the agarose. This step might be easier to accomplish with gloves off.
4. Let dry for 30 minutes by covering with a lid of a plate or tip box (assuring lids are not touching the agar!). The idea is to make sure the pad is not too wet

and not too dry. If it is too wet, your cells would float in a drop of water. At the same time, you want to make sure the pad is not too dry, since cells need moisture to divide. Make multiple pads in case something goes wrong with one.

## 3.2 Mounting the sample

This protocol is the same as in the *E. coli* growth module.

1. Your TAs will provide you with two samples of cells: one at an appropriate concentration for our growth movies (about $10^6$ cells/mL), and another at a higher concentration (about 10-100x more concentrated). While not strictly necessary, the latter is used to make finding the correct focal plane easier (see below).
2. Cut three 3 to 5 mm squares of agarose. Cover the pads with a lid of a plate or a tip box lid while they are drying (without touching them!). The idea is to make sure the pad is not too wet and not too dry. If it is too wet, your cells would float in a drop of water. At the same time, you want to make sure the pad is not too dry, since cells need moisture to divide. (We're making multiple pads in case something goes wrong with one.)
3. Spot 1 µL of cells on each pad. Remember that you should NOT touch the pad with the pipette tip! Two pads should get cells from the low concentration and one pad from the high concentration cultures.
4. Cover the pads with a lid of a plate or a tip box lid and let them dry for a few minutes until the water spot is no longer visible. Depending on how dry the pad is, we might put it in the 37°C incubator to speed up drying.
5. Gently transfer the pads to a glass-bottomed dish as demonstrated by your TA. Remember that the side where you spotted the cells should face the bottom of the petri dish. Be sure to mark the edge of the dish with a Sharpie and note in your notebook the orientation of the pads in the dish so you know which is which when you go under the microscope (remember to account for the mirror image flip when inverting the dish!).
6. Wait 15-20 minutes after placing the sample on the microscope for the temperature to reach and stabilize at 37°C.

## 3.3 Acquiring the images

1. Using 100x oil objective, find the proper focal plane such that bacterial edges are properly defined. You should use brightfield illumination for phase contrast microscopy while doing this. The TAs can help you with this, and also setting up Köhler illumination.

2. While in focus, determine a good exposure time for the fluorescence channel that bacteria cells are visible.
3. Determine the frequency, channels, and exposure intensity (in each channel) for image acquisition. Take images accordingly, making sure to maintain focus. Make sure you record the times at which you acquired your images. You will need to include these in metadata that is shared with the class.

# 4 Assignment

**Problem 0** (Summary)**.**
Write a summary of this experiment, its objectives, and conclusions between one paragraph and one page in length.

**Problem 1** (Segmentation)**.**
**This problem is due a week earlier than the rest of the assignment.**Following the protocol for segmentation and naming from the Image Processing III tutorial, obtain the median intensity for a lineage of cells. If your last name starts with A through R, use the TreRL movie. Otherwise, use the LacI movie. Submit your segmentation as a single CSV file. You must strictly adhere to the naming convention and CSV file format laid out in that tutorial.

**Problem 2** (Analysis of periods)**.**
From the data provided to you from all experiments, make informative plots of the time courses of fluorescent intensities for both treRL and LacI oscillators. You can include many plots of fluorescent intensity on the same plot. Comment and any features you notice in your plots. Also give rough estimates of the oscillator periods of the treRL and LacI oscillators.

**Problem 3** (The necessity of enzymatic degradation)**.**
By numerically solving the mathematical model for this system, demonstrate that in the absence of enzymatic degradation ($\gamma_1 = 0$), the dilution/degradation rate must be more than an order of magnitude larger in order to get sustained oscillations. That is $\gamma_0$ must be more than ten times the value based on cellular growth rate. This is untenable, which is why active degradation is necessary. This is an example of how mathematical modeling can lead to insights of circuit design.

**Problem 4** (Design choices: Antibiotic resistance)**.**
Why do the plasmids have genes imparting different antibiotic resistance? And why do we grow up the cells in the presence of both Kanamycin and Carbenicillin?

**Problem 5** (Design choices: Promoters)**.**
Why is it a good idea to have the reporter under control of the same promoter as the repressor?

**Problem 6** (Design choices: Enzymatic degradation)**.**
Why is it a good idea to add a degradation tag to the sfYFP reporter so that it is also degraded by ClpXP?

**Problem 7** (Amplitudes are difficult)**.**
Why is it difficult to compare amplitudes of the oscillations across experiments?