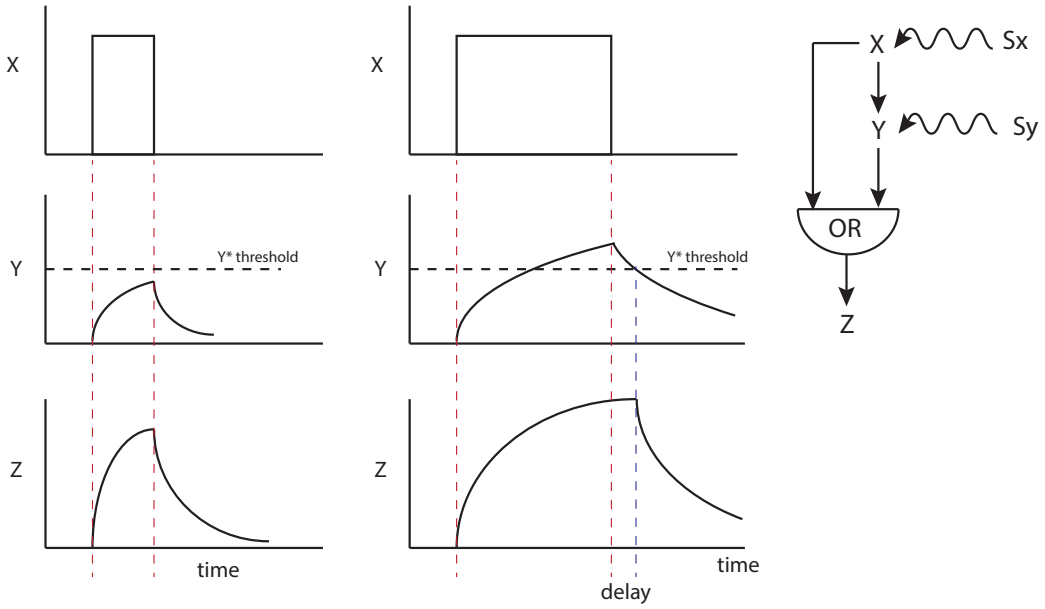
PLEASE DO NOT SPREAD

Homework 2 solutions
Paul Hudson
BB1110 2015

1A)



Here we assume that a pulse of Sx instantaneously activates X (a good assumption if Sx allosterically activates X, like cAMP and CAP). Thus, a pulse of Sx is a pulse of X. We will also assume here that Sy is present- like arabinose or lactose.

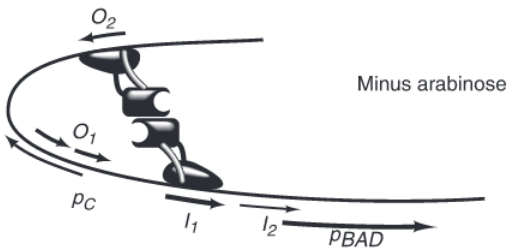
In the first case, a pulse of Sx is causes activation of X and activation of Y. Because Z is under an OR gate, it is activated even if Y has not yet accumulated to the Y* threshold. Removal of X causes a decrease in Y and an *immediate* decrease in Z. This is because, even though Y is still present, it is not about the threshold Y* to trigger Z.

In the second case, we have a longer pulse of Sx. Here X is activated and Y is also activated, where it accumulates over the threshold Y*. When X is removed, Z does not immediately decay as before. This is because Y is still over Y* for some time. Once Y decreases below Y*, Z begins to decay. So we have a delay in turning OFF Z, whereas in the feed forward AND gate, we had a delay turning ON Z.

1 B) The arabinose catabolism operon is the most well known example of both feed-forward regulation and an AND gate. Here are some pictures of what the region around the arabinose locus on E.coli genome looks like. Pc is the promoter for *araC*

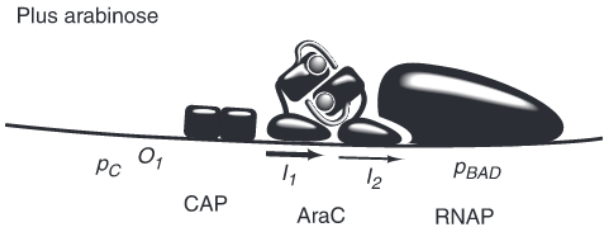
and Pbad is promoter for the *araBAD* operon, which are genes that degrade arabinose. Note that Pc is read in the opposite direction as Pbad.

In the first picture, there is no arabinose. You can see the AraC protein forms a dimer and causes a DNA loop, which blocks transcription of Pbad as well as its own promoter Pc. It is a repressor in this condition.



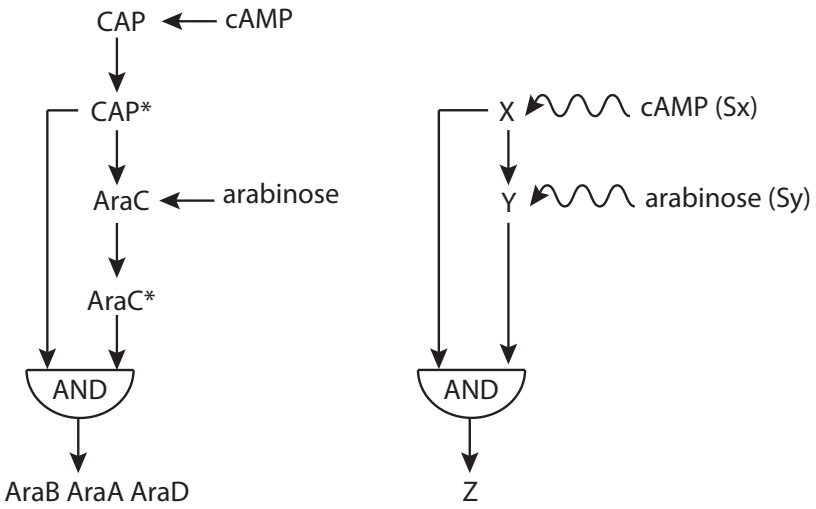
The case when +cAMP and –Arabinose. No transcription of Pbad. No transcription of Pc.

The CAP-binding site is for both promoters Pc and Pbad so that a cAMP signal is needed to transcribe Pc and thus get AraC. When arabinose is also present, AraC undergoes a conformational change to AraC* and binds to the I1 and I2 sites. This new arrangement is favorable for RNA polymerase to bind to the Pbad promoter. Note that CAP is also needed for activation of Pbad.



The case when +cAMP and +Arabinose. Yes transcription of Pbad. Yes transcription of Pc.

Figure from Robert Schlief FEMS Microbiology Reviews 2006.



Left: A detailed diagram of the arabinose "AND" gate. CAP* is CAP that has been activated by cAMP. Otherwise it cannot bind DNA. AraC* is AraC that has been activated by arabinose, and AraC* is an activator.

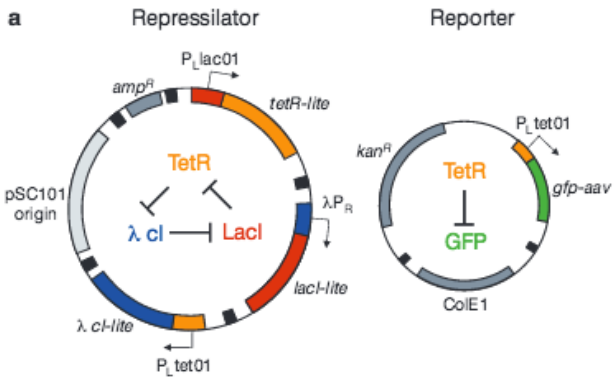
Right: A simplified version of the arabinose AND gate. We assume that CAP → CAP* is instantaneous in the presence of cAMP (a good assumption for allosteric activation of a protein by a small molecule). We assume that AraC → AraC* is instantaneous in presence of arabinose (a good assumption for allosteric activation of a protein by a small molecule).

In part A) we assumed that arabinose was already present, and a pulse of "cAMP" occurred.

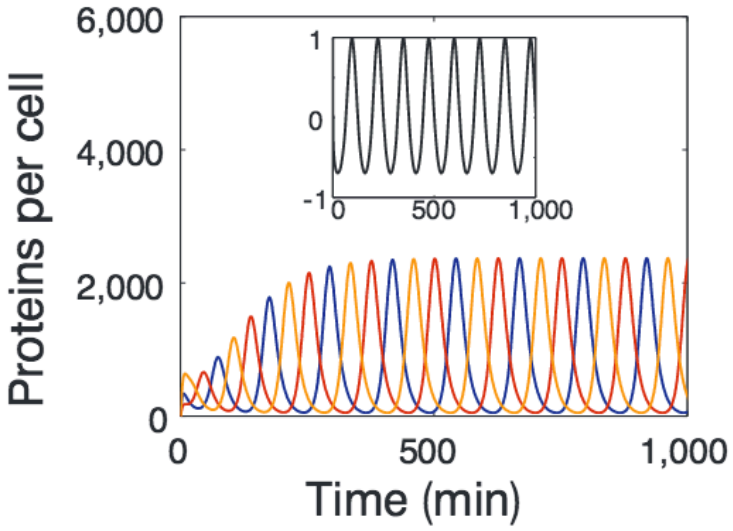
For an OR gate, you can imagine that a double promoter where the DNA looping is sensitive to two proteins, AraC and LacI. Here you could design the binding sites such that looping (and thus RNA polymerase occlusion) only happens if both AraC and LacI are present. That is, AraC nor LacI alone can hold the loop. So as soon as one is activated, the loop falls apart.

You can also see that design of DNA binding sites (or protein amino acids!) can change how strong the effect is.

2. This is the repressilator. It was first devised in 2000 (see Eilowitz and Leiber 2000 Nature for more).



Numerical solution to the Repressilator



Where TetR is yellow, CI is blue (GFP would follow CI), and LacI is red.

Several assumptions are made about the Represillator as solved here.
First, that proteins degrade quickly, such that when mRNA is gone they are quickly depleted. Also, that binding interactions for each repressor with its cognate promoter is the same (i.e. same Km for repressor::promoter binding for each repressor, where Km is the number of repressor protein needed to reduce transcription 1/2).

Oscillatory regulation is most prominently seen in circadian clocks. It is also present as a way to “synchronise” the output of several clocks, as was shown in a later paper, for those interested see Garcia-Oljavo et al PNAS 2004.

Extra stuff:
The mRNA and protein concentrations of for each repressor can be modeled over time as six differential equations, 3 for mRNA and 3 for protein.

$$\frac{dm_i}{dt} = -m_i + \frac{\alpha}{(1+p_j^n)} + \alpha_0 \qquad \begin{pmatrix} i = lacI, tetR, cI \\ j = cI, lacI, tetR \end{pmatrix}$$

$$\frac{dp_i}{dt} = -\beta(p_i - m_i)$$

Note that mRNA production rate (dm/dt) is dependent on a protein concentration (this is the repressor, and here p is actually normalized by Km of the repressor)

The protein production rate (dp/dt) is dependent on the amount of mRNA as well as the amount of protein (degradation is a 1st order process).

In the solution above, they solved these equations numerically using following:
Promoter strength: 0.0005 (repressed)
Promoter strength: 0.5 (induced)
Average translation efficiency: 20 proteins per mRNA transcript
Protein half life: 10 min
mRNA half life: 2 min
Km: 40 (repressor molecules per cell)

3. Yeast cannot synthesize NADPH or ATP in order to fix the CO2 unless glucose is present. This is typically done via photosystems in plants (chloroplasts)

4. Glycerol production is predicted to be zero in the mutant. It seems that glycerol (a 3- carbon molecule) was an "NADH sink" in the wild-type. When CO2 is present to be fixed, it changes the stoichiometry so that glycerol cannot be the NADH sink. Instead, the 2-carbon Ethanol becomes the sink. This somewhat unpredictable result is in fact stoichiometrically the best way to dump those electrons (while still growing) is to fix CO2 instead.

5. Metabolic model is undetermined because there are many more reactions than metabolites. So there are many possible distributions of metabolite fluxes with satisfy mass balances. The objective function gives the system a "goal" to reach, so that the possible solutions is narrowed. The most common objective function is the maximize flux toward the "biomass" reaction (represents growth).

6. The model allows you to set constraints on the flux solution for each reaction individually. A "knockout" would mean fixing that particular flux to zero, and resolving the model (keep the growth objective function). You can then see how the other reaction rates would change when that one is gone. Alternatively, algorithms can stepwise set every reaction to zero and resolve the model, and identify which knockouts increased flux of a particular reaction (say, Ethanol). This way you could find which knockouts (or combination) would increase ethanol. Computationally expensive, however.

7. First sequence the genome. Predict open reading frames using known motifs such as promoters, terminators, even look for possible ORFs in all reading frames. Then, BLAST these ORFs against known sequences in a database. This will give you probably enzymatic function and thus what reactions are present. Then you would need to manually go in and add known side reactions.

8. UV mutagenesis (see Lecture 1!), chemical mutagenesis (here use a molecule which binds to DNA so that when DNA is replicated there is an error). Alternatively could do error-prone PCR on a particular gene and re-insert mutant of this into genome. Alternatively, just growth for a long time means that mutations will occur randomly and perhaps accumulate in the genome. We also discussed a new discovery that, under stress conditions (such as heat, or an antibiotic like ampicillin), E. coli activates an "error-prone" polymerase, which appears to have the purpose of purposefully making mutations. This is like a defense mechanism for a community… the mutations could be lethal to some members, but if a few get a mutation that allows them to survive the stress (heat, antibiotic), they will continue the species.

9. Important that you know difference between screen and selection. In our screen, we could grow the E. coli on a plate containing glucose and an oil component (like fatty acid). Perhaps the fatty acid has a chemical dye attached, so that it is colored.

When the E. coli eats it, the color around that colony disappears, and creates a "halo." Your plate will have many colonies with no halo, these are cells that do not consume the oil compound. They survive, however, because of the glucose. Those with big halos after a few days ate the oil compound the fastest.

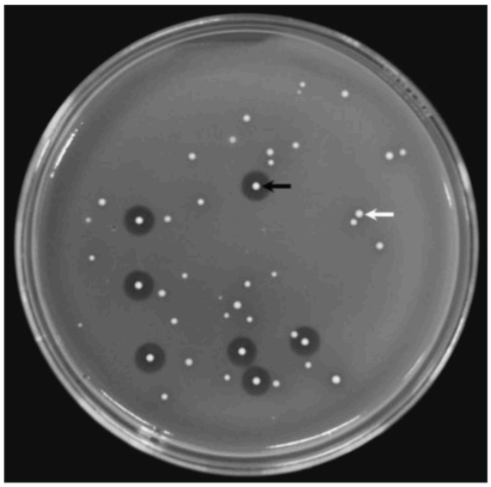


FIG. 1. Screening of bacteriocin-producing LAB. The black and white arrows indicate examples of positive and negative colony with or without a clear halo zone, respectively.

Here is an example of screening LAB for production of a toxin (Sunwanjida et al 2007). In this example, the bottom side of the agar is covered with "tester" bacteria. When t these grow, it gives the plate a hazy look. The topside of the agar is covered with a library of "LAB" bacteria. These LAB bacteria may or may not produce a toxin. As the LAB cells grow, the toxin diffuses out of the colonies, diffuses through the agar, and kills the "tester" bacteria. Where "tester" bacteria do not grow, there is a clear halo on the agar plate. The LAB colonies at the center of these halos produce toxin.

10. For selection, you can imagine plating the cells on a plate which contains *only* oil component. That way, they have to eat it or they die. This is a *selection*