Principles of Synthetic Biology Midterm

October 30, 2012

Instructions: This exam is an open book / open notes exam. You are not permitted to use any electronic devices during the course of this exam. You will have 90 minutes to complete the test. Please use exam booklets for your solutions. Answers included elsewhere will not be graded.

This exam is a bit lengthy so please exercise sound time management. Good Luck!

1. Circuit Parts and Compositors (10 Points)

Suppose we transformed the following transcriptional regulatory network into a cell:

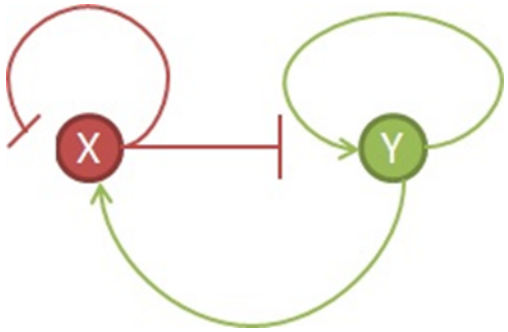


Fig. 1: A model transcriptional regulatory network

1.1. Qualitatively describe one or more possible behaviors of this system. Under what conditions would these behaviors manifest themselves? Furthermore, what dependencies on the host cell exist?

1.2. List the corresponding *PARTS* of this system and provide one possible mathematical representation of these parts. Draw a boundary around the respective parts on figure 1 itself.

1.3. List the corresponding *COMPOSITORS* of this system and provide one possible mathematical representation of these compositors.

2. Circuit Analysis (30 Points)

Consider the circuit in figure 2. In this system, *A* is a transcriptional activator of gene *O* that binds to *O*'s promoter in a non-cooperative, Michaelis-Menton fashion. *B*, whose production is induced by the concentration of Arabinose presented, inactivates *A* by binding to it and forming a heterodimer complex *AB* that is no longer able to activate *O*. Assume that there is some basal production of *O* even when no activator is bound to its promoter. For simplicity, assume that all binding and unbinding reaction achieve equilibrium rapidly. Use the following sub-questions to guide you in this problem.

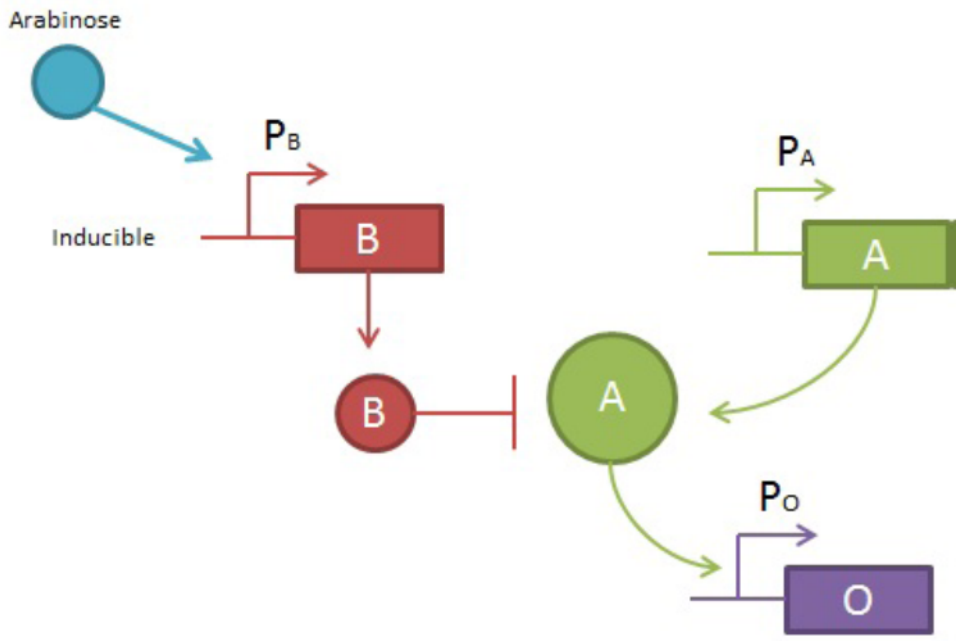
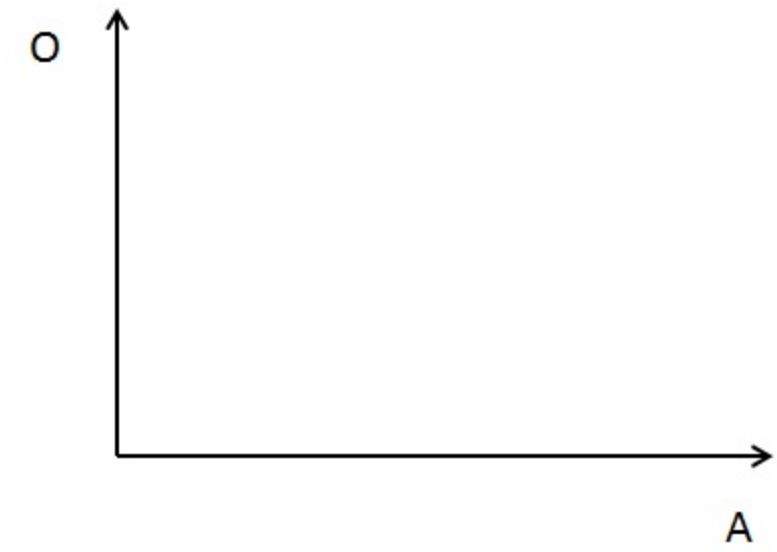


Fig. 2: A model gene circuit

2.1. If the formation of the heterodimer complex *AB* is vastly favored over free *A*, what would you expect to happen at *low* induction, in which the total amount of *B* in the cell is much less than that of *A*? What would you expect to happen at *high* induction?

2.2. Qualitatively describe the effect of increasing the total amount of *B* in this system. Sketch the output-input (*O* vs. *A*) curve in the graph below as you turn on induction of *B* with arabinose. Speculate on any changes to the steady state levels of *O*. Sketch three curves representing increasing induction and clearly label them as either *low*, *medium*, or *high* induction.



2.3. Qualitatively describe what happens in the regime where the total concentrations of *A* and *B* in the system are close to each other.

2.4. Now consider the same scenario with a positive feedback loop in which *A* can catalyze its own production (see figure 3). Suppose you transformed this system into *E. coli*. Bacteria with high Yellow Fluorescent Protein (YFP) expression (above some basal threshold) are considered ON, whereas those that do not are considered OFF.

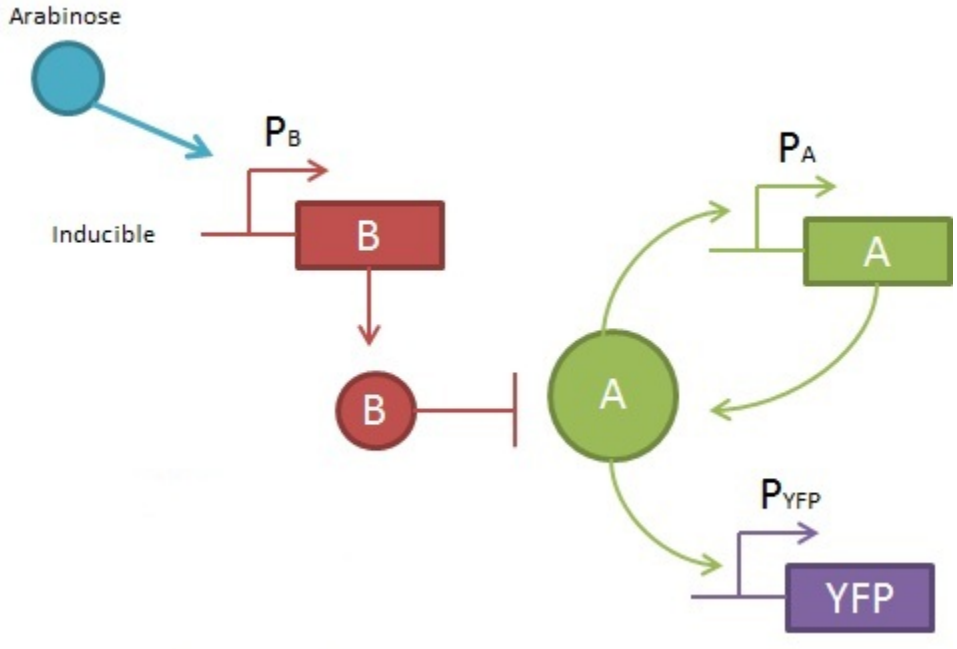


Fig 3: **A model gene circuit with a positive feedback element**

2.5. Speculate as to what effect(s) the positive feedback loop might have on circuit dynamics as opposed to the open loop case (figure 2). Consider the following experiment: Grow bacteria in arabinose deficient media and then inoculate them into media containing intermediate concentrations of arabinose. Measure YFP levels for each cell. Separately, grow bacteria at full induction of *B* for the same length of time and then inoculate a sample into the same intermediate concentrations of arabinose. Measure YFP levels for each cell. How would you expect these two sets of measurements to compare? Would they be the same or different and why? What useful property does this circuit have?

3. Cell Classifier (30 Points)

You find three markers *A, B,* and *C* which you can use to identify a cell line. You would like to build a circuit that implements the appropriate logic (given below) and only produces GFP during specific instances of marker expression that match this profile.

A	*B*	*C*	**G (GFP)**
0	0	0	**1**
1	0	0	**0**
0	1	0	**1**
1	1	0	**0**
0	0	1	**0**
1	0	1	**0**
0	1	1	**1**
1	1	1	**0**

3.1. Find the simplest logic expression, $G = G(A, B, C)$, that captures the input-output relationship demonstrated in the truth table above. Write this as a logic function.

3.2. Draw the simplest logic circuit, using as few logic gates as possible, that captures the input-output relationship demonstrated in the truth table above. You may only use NOT, 2-input AND, and 2-input OR logic gates.

3.3. List any hazards in your circuit. Again using as few logic gates as possible, draw the simplest hazard-free circuit. If your design does not have any hazards, write NO HAZARDS, and proceed to the next question.

3.4. Implement your hazard free circuit using only the following biological parts: Repressor protein (Rep), protein comprising an Activation domain (ActD) coupled with a Leucine zipper (Lx), and a protein comprising a DNA binding domain (DBD) coupled with a Leucine zipper (Ly). The corresponding singly-activated and singly-repressed promoters are also available. For this question, assume that Leucine zippers can form heterodimers specifically. In your answer, draw the biological circuit, and indicate the specific Leucine zipper pairs (i.e L*i* binds to L*j*).

4. Noise Generator (30 Points)

Galenman is interested in understanding bacterial population dynamics. He proposes to construct a gene circuit that can drive a homogeneous bacterial population into two distinct subpopulations in such a manner that the ratio of the two subpopulations can be controlled by a chemical. To accomplish this feat, he intends to design a noise generator whose noise level can be modulated by an exogenous chemical *C*, and then connect this noise generator to a bi-stable switch. The population type is determined by whether or not the output of the bi-stable switch, which is based on a positive feedback loop, is high. Galenman finds a tetrameric activator *A* and promoter *P* pair. The promoter can be turned off by adding chemical *B*. He also finds a series of degradation tags that he can use to modulate protein concentration in the cell. The protein production curve from this promoter is shown below (figure 4).

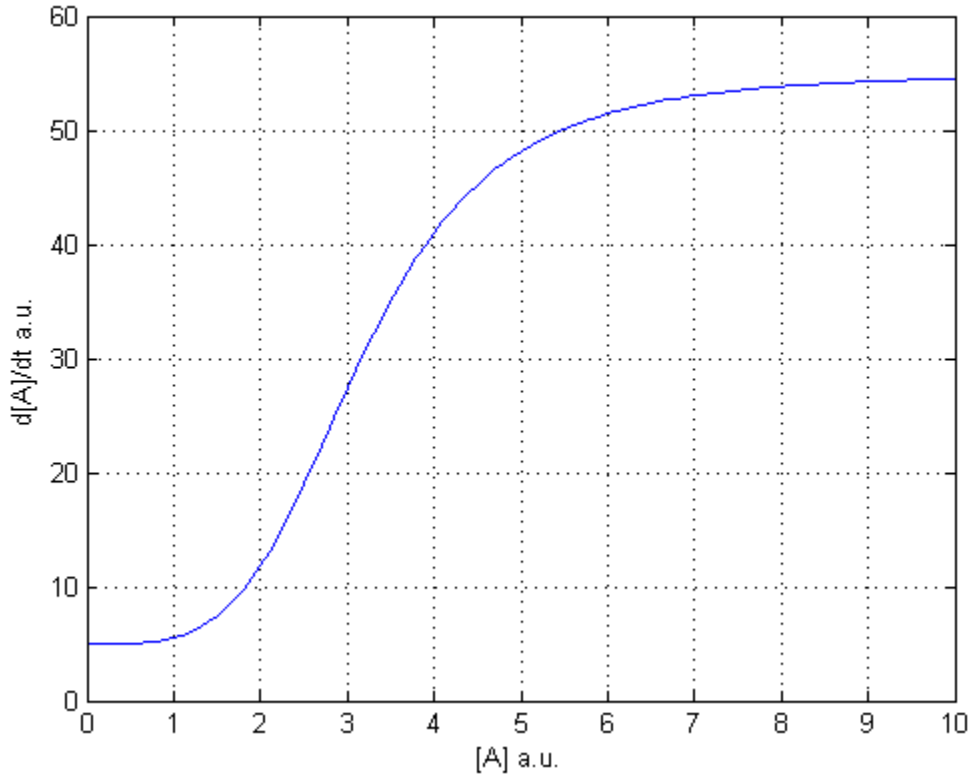


Fig. 4: Production curve of A from P

4.1. From the table below, choose a degradation tag that can form a bi-stable switch. On figure 4 itself:

Tag Name	Degradation rate (a.u. per unit time)
Slowtag	3
Medtag	5
Fastag	9
Spedtag	12

4.1.1. Sketch the degradation curve for the chosen degradation tag

4.1.2. Circle any and all stationary points

4.1.3. Label the stability of each stationary point as either stable or unstable

4.1.4. Assuming that the tetramer formation is highly cooperative and fast relative to the transcription and translation of *A*, estimate the numerical value of the three parameters in the ODE provided below such that the model accurately fits the production curve in figure 4.

$$\frac{d[A]}{dt} = a + b\frac{k \cdot [A]^4}{1 + k \cdot [A]^4}$$

4.2. To make the noise generator, Galenman finds a low noise promoter that can reliably maintain its mRNA copy level at *m* copies per cell. Assuming that the translation rate, k_{tr} (events per unit time), is given by $k_{tr} = 2C$, where *C* is the concentration of chemical *C* and that translation events occur independently, compute the mean and variance of protein production per unit time in terms of *m* and *C*. Justify your result.

4.3. The output from the noise generator will be the same tetrameric activator. In order to control the population ratio, Galenman needs to determine the appropriate amount of *C*. To simplify the problem, assume that the distribution you computed in question 3.2 can be approximated by a continuous triangle function that peaks at the mean and drops to zero at one standard deviation of the mean, i.e at $ean \pm \sqrt{variance}$. Assume that both the bi-stable switch and the degradation processes are strictly deterministic and that the noise generator alone is stochastic. Calculate the concentration of *C* needed in order to set the *A-HIGH* population to 20% of the total population when $= 18$. Suggest an experimental procedure that could be used to recreate this scenario.