1. SYNTHETIC TRANSCRIPTIONAL CASCADES (25PT)

Cascades are ubiquitous regulatory motifs in cells. To study the functions of cascading motifs and their functions, you are about to construct synthetic cascades in bacteria cells.

First, you construct a cascade that is composed of two activators as shown in Figure 1. Protein A_1 is initially present in the cell in its inactive form. The input signal X_1 is added at time t = 0. Depending on the concentration of X_1, certain amount of A_1 ***rapidly*** becomes active and binds to the promoter of A_2, so that protein A_2 starts to be produced at rate β. When A2 levels exceed a threshold K_{A2}, GFP begins to be produced at γ.

All assume all proteins don't degrade, and they only have a dilution rate of α when the cell divides.

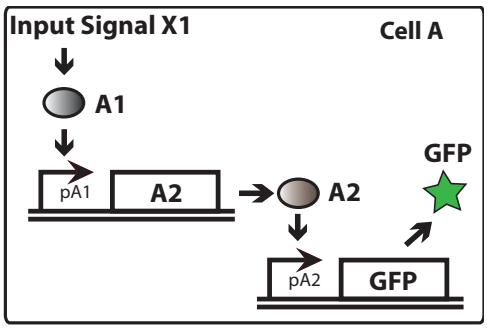


Figure 1

(A) After you constructed the circuit, you added 1uM of input signal molecule X1. However, you didn't observe any GFP signal after a long time. So you characterized the transfer curves for promoters pA1 and pA2, and the induction relationship between X1 and A1 (all axes are in logarithm10 scale). (5PT)

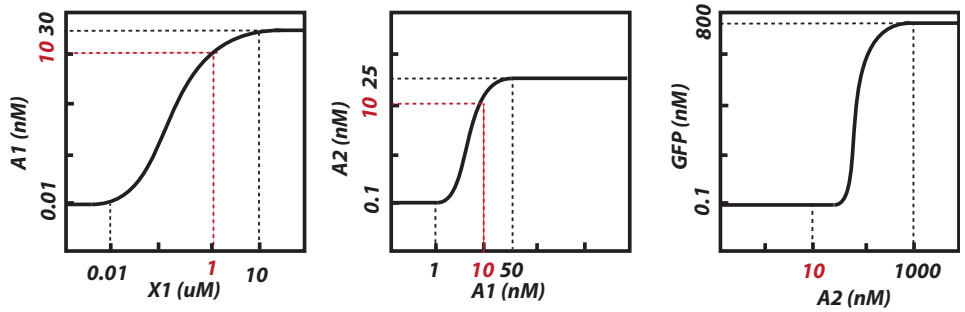


Figure 2

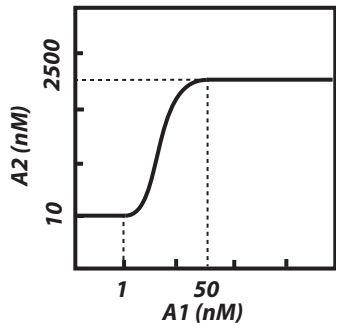
Explain why you didn't observe GFP signal.

As shown in the figure, with 1uM of input signal, we expect 10nM of A1, which will produce 10nM of A2. 10nM of A2 is not enough to turn the promoter ON.

(B) Your advisor points to you that these transfer curves are potentially limiting the GFP signal you might observe. Identify the problem, and propose how you might improve the circuit. (5PT)

There is a signal-matching problem. The dynamic range for A1 is 0.01nM to 30nM, corresponding to the dynamic range of A2 as of 0.1nM to slightly less than 25nM. However, even at the maximal amount of 25nM, this amount of A2 couldn't induce GFP.

One solution is to increase the strength of RBS for A2 by a factor of around 100-fold so the transfer curve shifts up.



(C) After improving the circuit, you are finally able to observe GFP signal after you add the input signal X1. The time delay of the circuit τ, is defined as the duration between the addition of the input signal (t = 0) and initial expression of GFP (t = τ). Explain briefly the source of the time delay in this circuit, and calculate how much is the delay using the given parameters (5PT).

The delay is caused by the production time for A2 to above the threshold K1.

It is given that the production rate of A2 is β, and its dilution rate is α.

$$\frac{d[A_2]}{dt} = \beta - \alpha[A_2]$$

Solve this equation, gives

$$[A_2(t)] = \frac{\beta}{\alpha} - C \cdot e^{-\alpha t}$$

At t = 0, [A2] = 0, so

$$[A_2(t)] = \frac{\beta}{\alpha}(1 - e^{-\alpha t})$$

Let

$$[A_2(\tau)] = K_1$$

we obtain

$$\tau = \frac{1}{\alpha}\ln(\frac{\beta}{\beta - \alpha K_1})$$

Then you set to construct a second cascade that is composed of two repressors as shown in Figure 3. Protein R_1 is initially present in the cell in its inactive form. The input signal Y_1 is added at time t = 0. As a result, R_1 rapidly becomes active and binds to the promoter of R_2, so that expression of R_2 is shut off. Assume the initial concentration of R_2 is η. When R_2 drops below a threshold of K_{R2}, GFP begins to be produced at rate γ.

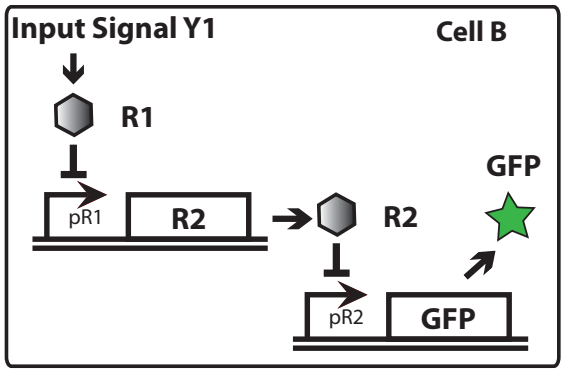


Figure 3

(D) Calculate the time delay using the given parameters. Explain the source of the time delay in this circuit? (5PT)

The delay is caused by the degradation/dilution time of R2. Because dilution is caused by cell replication, so this is inherently limited by cell growth rate.

$$\frac{d[A_2]}{dt} = -\alpha[A_2]$$

Solve this equation, gives

$$[A_2(t)] = C \cdot e^{-\alpha t}$$

At t = 0, [A2] = η, so

$$[A_2(t)] = \eta e^{-\alpha t}$$

Let

$$[A_2(\tau)] = K_2$$

we obtain

$$\tau = \frac{1}{\alpha}\ln(\frac{\eta}{K_2})$$

(E) A good approximation for repressible promoter activity is

$$f(X) = \frac{\varphi}{1+(\frac{X}{\kappa})^n}$$

Where X is the concentration of repressor protein, and $f(X)$ is the promoter activity. Briefly explain the physical meanings of the parameters φ, κ and n, and the assumption(s) making this expression valid. (5PT)

Assumptions: (1) Fast kinetics of multimerization; (2) Abundant repressor;

φ - maximum strength of the promoter (without repressor)

κ - the amount of repressors need for 50% repression also reflects the concentration requirements of repressors.

n – cooperatitivity, reflect how sigmoidal the transfer curve would be.

(F) Both R1 and R2 bind to their promoters in the monomer form. Do you expect the steady-state transfer curve of the cascade circuit in Cell B will exhibit a higher sensitivity than that of the single repressor-promoter circuit? Write equations to explain. (5PT)

No, we don't expect the sensititivty would increase. The steady state transfer curve for single repressor promoter circuit is:

$$GFP(R_1) = \frac{1}{\alpha} \cdot \frac{\varphi}{1 + \frac{R_1}{\kappa}}$$

The steady state transfer curve for the cascade is a compound function:

$$GFP(R_2) = \frac{1}{\alpha_1} \cdot \frac{\varphi_1}{1 + \frac{R_2}{\kappa_1}}$$

$$R_2(R_1) = \frac{1}{\alpha_2} \cdot \frac{\varphi_2}{1 + \frac{R_1}{\kappa_2}}$$

So,

$$GFP(R_1) = \frac{1}{\alpha_1} \cdot \frac{\varphi_1}{1 + \frac{R_2}{\kappa_1}}$$

A little bit of math shows

$$GFP(R_1) = A - \frac{B}{1 + \frac{R_1}{C}}$$

Where $A = \varphi_1 / \alpha_1, B = \varphi_2 / (\varphi_2 + \alpha_2 \kappa_1), C = \kappa_2 (1 + \varphi_2 / \alpha_2 \kappa_1)$ are all constants. So the cooperativity of the cascade doesn't increase.

2. NEGATIVE FEEDBACK (25PT)

Feedbacks are important network motifs in biological systems. To understand their functions deeper, you will build and model a synthetic feedback circuit. You plan to use one transcription repressor and its promoter to build a simple circuit, so expression of the repressor protein will repress itself by repressing the promoter as shown in Figure 4.

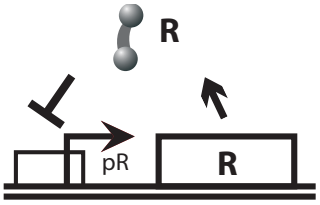


Figure 4

(A) Write down the chemical reactions for the circuit. (5PT)

$R + pR \leftrightarrow RpR$

$pR \rightarrow pR + mRNA$

$mRNA \rightarrow mRNA + R$

$mRNA \rightarrow$

$R \rightarrow$

(B) Comparing the negative feedback circuit to the simple circuits where there is no repression on the promoter, choose the correct answers.

- The negative feedback has (a) speed to reach its steady state;

a. faster

b. slower

c. same

- The negative feedback is (a) to the fluctuations in gene expression;

a. more robust

b. more susceptible

c. similar

(C) Assume the mRNA dynamics is already at its steady state. Also assume the promoter activity follows the step function:

$$f(R) = \varphi, \text{ if } R < \kappa$$

$$f(R) = 0, \text{ if } R \geq \kappa$$

φ and κ are all constants. Write down the ODE equations to describe the expression of R (5PT).

Because the mRNA is at its steady state, so $[mRNA] = f(R) / D_m$.
The expression of R is

$\frac{d[R]}{dt} = \alpha \cdot [mRNA] - \gamma[R]$ *becomes*

$$\frac{d[R]}{dt} = \alpha\varphi - \gamma[R], \text{ if } [R] < \kappa$$

$$\frac{d[R]}{dt} = -\gamma[R], \text{ if } [R] \geq \kappa$$

(D) You test the feedback circuit, and find that after a few cycles of oscillation, the circuit reaches its steady state concentration that equals to κ . Explain why there are oscillations and why the steady state concentration is κ .

When $[R] < \kappa$, *because* φ *is large, d[R]/dt>0, and the production of R will continue; and when* $[R] \geq \kappa$, *d[R]/dt<0 , and R is degraded. So when there are delays in the circuit such as the degradation rate is small, the circuit will exhibit small oscillations. But it finally reaches its steady state at* κ .

3. POSITIVE FEEDBACK (20PT)

Now consider a positive feedback circuit as shown in Figure 5, where a promoter drives the production of one activator protein A to activate the promoter. A binds to the promoter in the monomer form.

(A) Assume mRNA is always at steady state, and follow a similar promoter expression from Problem 1(E), write one equation to describe the system. (5PT) (*hint: you need an activated promoter expression*)

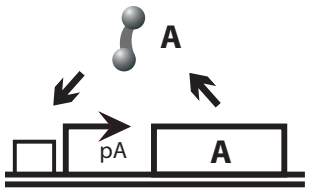


Figure 2

The promoter function in this case is

$$f(X) = \frac{\varphi(\frac{X}{\kappa})^n}{1+(\frac{X}{\kappa})^n},$$

Because A binds to the promoter in monomers,

$$f(X) = \frac{\varphi X}{\kappa + X}$$

mRNA is $[mRNA] = f(X)/D_m$

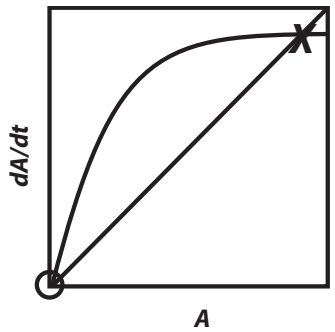
Then ,

$$\frac{dA}{dt} = \frac{\alpha\varphi}{Dm}\frac{X}{X+\kappa} - D_p X$$

(B) Your expression should look like

$$\frac{dA}{dt} = f(A) - g(A)$$

On a 2D plot where the x-axis is the amount of A, y-axis is the $\frac{dA}{dt}$, draw a curve of f(A) and a curve of g(A) as a function of A, and label the critical points. (5PT)



(C) How many steady states does this circuit have? Circle the stable steady states, and explain why it is stable. (5PT)

Only 1.

(D) Switches that possess two stable critical points, also called bistable switches, are especially interesting and useful in biology. Positive feedback loops are important for building bistable switches. To engineer the above circuit to be a bistable switch, you think of the following strategies:

1) Engineer the activator A so it only activates its promoter when it forms a dimer;

2) Change the protein degradation rate of activator A;

3) Tune the RBS to increase or decrease the translation rate of activator A;

4) Constitutively express another protein B that can sequester A from activating the promoter;

Identify which strategy(s) might allow you to engineer bistable switches, and explain why. (5PT)

Strategy of 1) and 4) would allow you to engineer bistable switches. 1) Increase the cooperative and makes the curve more sigmoidal. 4) shifts the curve to the right side and allows the two curve to have three crossing points.

STOCHASTIC SYSTEMS ANALYSIS

Biological systems are noisy and noise has been shown to be a critical driving force of various biological events. ODE simulations cannot adequately capture this noise.

a) Briefly explain why this is true and give an example.

Gillespie's stochastic systems analysis (SSA) simulates the biological system by treating reactions as probabilistic events. Let us consider a closed system with the following reaction dynamics.

$$2S_1 \xrightarrow{0.1/sec} 2S_2$$

$$S_1 + S_2 \xrightarrow{0.4/sec} 2S_1$$

$$S_2 \xrightarrow{0.6/sec} S_3$$

b) Using SSA, if the initial state is defined as $[S_1] = 1$, $[S_2] = 2$, $[S_3] = 0$, what reaction is most likely to occur for the first reaction step?

c) Keeping the same initial state, what is the probability that after two reaction steps, $[S_3] = 0$?

d) Now, let us redefine the initial state to be $[S_1] = 1$, $[S_2] = 1$, $[S_3] = 0$; again, what is the probability that after two reaction steps, $[S_3] = 0$?

e) After how long can we be certain that the system as defined in (d) is stable?

f) If you are designing a circuit and you would like to take advantage of biological noise, would you choose positive feedback or negative feedback? Intuitively explain your choice.

SOLUTION

a) ODE simulations cannot adequately capture this noise. Briefly explain why this is true and give an example.

The accuracy of deterministic analysis is heavily dependent upon reactions having repeatable, predictable behavior within the time frame of interest. Common assumptions in deterministic analysis are abundance and time progression to steady-state giving higher accuracy when applied to large-scale or long-term, continuous-time system dynamics. However in low molecular count, small-scale or short-term systems, the SNR is much lower, time becomes discretized, and noise plays a more significant role.

For example, when there exist only 5 molecules of ATP, but 10 are required to activate transcription, stochastic analysis will indicate this process will not occur. However, deterministic analysis is incognizant of species count and would not take this into account.

b) If the initial state is defined as $[S_1] = 1$, $[S_2] = 2$, $[S_3] = 0$, what is the probability of each reaction occurring for the first reaction step?

$$\cancel{P(R1) \propto \left\{k_1 * [S_1]^2 = \frac{0.1}{sec} * 1^2\right\} \propto 0.1}$$

$$P(R2) \propto \left\{k_2 * [S_1] * [S_2] = \frac{0.4}{sec} * 1 * 2\right\} \propto 0.8$$

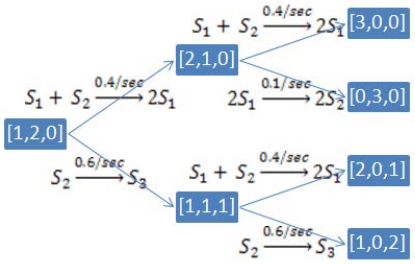
$$P(R3) \propto \left\{k_3 * [S_2] = \frac{0.6}{sec} * 2\right\} \propto 1.2$$

$$P(R2) + P(R3) = 1$$

$$\rightarrow P(R1) = 0, P(R2) = \frac{0.8}{1.2+0.8} = 0.4, P(R3) = \frac{1.2}{1.2+0.8} = 0.6$$

c) Using the same initial conditions as (a) what is the probability that after two reaction steps, $[S_3] = 0$?

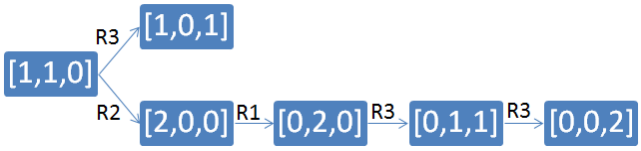
ALL 2-step trajectories:



$S_3 = 0$ for the trajectories {->R2 ->R2} and {->R2->R1}. Because this encompasses all possible trajectories if R2 occurs as the first reaction, P({->R2 ->R2}) + P({->R2->R1}) = P({->R2}).

$$\text{p(R2,[1,2,0])} = \frac{P(R2)}{\cancel{P(R1)+}P(R2)+P(R3)} = \frac{k_2*[S_1]*[S_2]}{k_2*[S_1]*[S_2]+k_3*[S_2]} = \frac{2k_2}{2k_2+2k_3} = 0.4$$

d) Now, let us redefine the initial state to be $[S_1] = 1$, $[S_2] = 1$, $[S_3] = 0$, what is the probability that after two reaction steps, $[S_3] = 0$?



{p(R2,[1,1,0]->[2,0,0])}*~~{p(R1,[2,0,0]->[0,2,0]) =~~ {p(R2,[1,1,0]->[2,0,0])}

$$\frac{P(R2)}{\cancel{P(R1) +} P(R2) + P(R3)} = \frac{k_2 * [S_1] * [S_2]}{k_2 * [S_1] * [S_2] + k_3 * [S_2]} = \frac{k_2}{k_2 + k_3} = 0.4$$

e) After how long can we be certain that the system with the same initial conditions as (d) is stable?

There are only two possible system trajectories:

(1) ->R3
(2) ->R2->R1->R3->R3
   a. k2 = 0.4/sec -> 2.5 seconds
   b. k1 = 0.1/sec -> 10 seconds
   c. k3 = 0.6/sec ->1.67 seconds
   d. k3 = 0.6/sec ->1.67 seconds
   e. TOTAL = **15.83 seconds**

f) If you are designing a circuit and you would like to take advantage of biological noise, would you choose positive feedback or negative feedback? Intuitively explain your choice.

I would choose negative feedback which would ensure the system remained at low molecular concentrations, retaining low SNR. Positive feedback would tend towards stability at much higher concentrations, making deterministic approaches more applicable and noise less significant.

LOGICAL REDUCTION

You are tasked with designing a biological reporter which monitors two separate threshold concentration sensors (S1 and S2) for two different biological species. Each sensor is comprised of two concentration level detectors (a and b), one set to activate at a lower threshold, and one set to activate at a higher threshold concentration, respectively. This enables concentration segmentation into 3 separate regions: (1) low – both a and b are inactive, (2) middle – only a is active, and (3) high – both a and b are active. Your biological reporter should take in the output signals S1a, S1b, S2a, and S2b and produce GFP if neither sensor reports concentrations in the high region, but at least one sensor reports in the middle region.

a. Derive a simplified logic expression for the system f(S1a, S1b, S2a, S2b) = GFP.
b. Sketch a corresponding logic circuit using the least number of total gates. Multi-input gates are allowed.
c. Reduce this to an optimized number of AND, OR, and NOT gates such that if each gate has a delay of 1, the time from change in input to change in output is minimized.
d. You are asked to additionally incorporate the feature of "blinking" fluorescence if either S1 or S2 reports concentration in the low region. Describe what design features you would need to implement this.

SOLUTION

LOGICAL REDUCTION

You are tasked with designing a biological reporter which monitors two separate threshold concentration sensors (S1 and S2) for two different biological species. Each sensor is comprised of two concentration level detectors (a and b), one set to activate at a lower threshold, and one set to activate at a higher threshold concentration, respectively. This enables concentration segmentation into 3 separate regions: (1) low – both a and b are inactive, (2) middle – only a is active, and (3) high – both a and b are active. Your biological reporter should take in the output signals S1a, S1b, S2a, and S2b and produce GFP if neither sensor reports concentrations in the high region, but at least one sensor reports in the middle region.

a. Derive a simplified logic expression for the system f(S1a, S1b, S2a, S2b) = GFP.

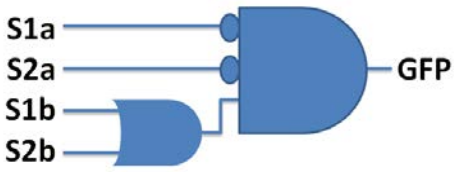
S1a	S1b	S2a	S2b	GFP
0	0	0	0	**0**
0	0	0	1	**1**
0	0	1	1	**0**
0	1	0	0	**1**
0	1	0	1	**1**
0	1	1	1	**0**
1	1	0	0	**0**
1	1	0	1	**0**
1	1	1	1	**0**

The shaded states are unreachable and may be set to any desired logic for maximum reduction flexibility. In this case, they don't lend any benefit, so a low response is preferable to high to avoid any potential transient errors.

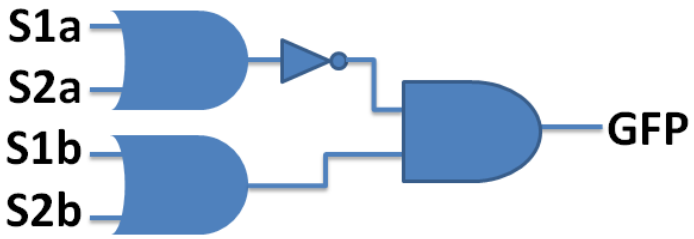
S1a,b→ S2a,b↓	00	01	11	10
00	0	1	0	0
01	1	1	0	0
11	0	0	0	0
10	0	0	0	0

S1a'S2a'S2b + S2a'S1a'S1b = S1a'S2a'(S1b + S2b).

b. Sketch a corresponding logic circuit using the least number of total gates. Multi-input gates are allowed.



c. Reduce this to an optimized number of AND, OR, and NOT gates such that if each gate has a delay of 1, the time from change in input to change in output is minimized.



d. You are asked to additionally incorporate the feature of "blinking" fluorescence if either S1 or S2 reports concentration in the low region. Describe what design features you would need to implement this.

A blinking feature may be implemented by the incoherent feed-forward motif ring oscillator activated when the output of S1a + S1b = 0, or S2a + S2b = 0.