Each homework consists of 3 problems, and you are expected to spend 30 min to 1 hour on each problem, but definitely less than 1 hour. If you find yourself spending more than 1 hour, you are probably overthinking about it. The optional problems may take significantly longer, so you can skip if you are short on time. But if you are interested in exploring further, the fun you get from working on the optional problems is definitely worth it!

1 Building blocks of diverse bioregulation

We learned about the three archetypal behaviors of bioregulation from one binding reaction, namely saturation, bottleneck, and ultrasensitivity. Biological systems can build diverse bioregulatory behaviors out of these fundamental elements of bioregulation. We investigate some of them here.

1.1 Competitive binding

Consider the following binding network.

$$G + R_1 \rightleftharpoons C_1, \;\; G + R_2 \rightleftharpoons C_2. \tag{1}$$

This is an abstract binding network model, but it may help by considering a specific biological context. We can view G as a gene, with repressors R_1 and R_2 binding with it competitively, so either R_1 binds to form a complex C_1 or R_2 binds to form a complex C_2, but not both. Let K_1 and K_2 denote the dissociation constants of the two binding reactions. The conserved quantities are the total concentration of gene molecules $G_{\text{tot}} = G + C_1 + C_2$, the total concentration of repressor one, $R_{1,\text{tot}} = R_1 + C_1$, and the total concentration of repressor two, $R_{2,\text{tot}} = R_2 + C_2$.

The activity species is G, since it is the unrepressed gene with transcriptional activity. The concentrations that can be regulated are those of the repressors. Therefore, described in terms of input-output, we have G as the output, $R_{1,\text{tot}}, R_{2,\text{tot}}$ as inputs, and G_{tot}, K_1, K_2 as parameters.

The idea of this competitive binding behavior is that the gene has only one binding site for it to be repressed, and we have two types of repressors, R_1 and R_2, that both represses G at this binding site. Therefore, R_1 and R_2 competes to bind with G. We would like to see how does the repression effect change with the amount of repressors R_1 and R_2, and how does this compare with just one repressor.

1. Consider the behavior of this binding network under the overabundance limit where $R_{1,\text{tot}}$ and $R_{2,\text{tot}}$ are overabundant, while G_{tot} is left variable. So this behavior's dominance condition is $R_{1,\text{tot}} \sim R_1$ and $R_{2,\text{tot}} \sim R_2$. Show that the following holds under this dominance condition,

$$G \sim G_{\text{tot}} \frac{1}{1 + \frac{R_{1,\text{tot}}}{K_1} + \frac{R_{2,\text{tot}}}{K_2}}. \tag{2}$$

   This is the saturation behavior with two saturating variables. Note that, if we view this as a function of $R_{1,\text{tot}}$ alone, i.e. considering $R_{2,\text{tot}}$ as a parameter as well, then we can compare what does this additional binding reaction do by comparing the above expression with the one-variable saturation function

$$G \sim G_{\text{tot}} \frac{1}{1 + \frac{R_{1,\text{tot}}}{K_1}}. \tag{3}$$

   We see that increasing $\frac{R_{2,\text{tot}}}{K_2}$ is effectively "flattening" the saturation curve of G in terms of $R_{1,\text{tot}}$.

   Draw the function of G vs $R_{1,\text{tot}}/K_1$ for different $R_{2,\text{tot}}/K_2$ to see this.

2. We can visualize this behavior. Since there are two inputs $R_{1,\text{tot}}$ and $R_{2,\text{tot}}$, we can consider a 2D graph with $R_{1,\text{tot}}$ as x-axis and $R_{2,\text{tot}}$ as y-axis. There are three dominance regimes included in this behavior, namely $G_{\text{tot}} \sim G$, $G_{\text{tot}} \sim C_1$ and $G_{\text{tot}} \sim C_2$. Show that these dominance conditions of G_{tot} for each regime corresponds to validity conditions involving $R_{1,\text{tot}}$, $R_{2,\text{tot}}$, K_1 and K_2. For example, the first dominance regime $G_{\text{tot}} \sim G$ should correspond to $1 \gg R_{1,\text{tot}}/K_1, R_{2,\text{tot}}/K_2$.

   Then, for each of these dominance regimes, draw their regions of validity on the 2D input graph. Compare this with your drawings of $G(R_{1,\text{tot}}/K_1)$ for different $R_{2,\text{tot}}/K_2$ in the previous problem.

3. What is the condition on the parameters G_{tot}, K_1, K_2 for this behavior to hold? Show that $R_{1,\text{tot}} \sim R_1$ and $R_{2,\text{tot}} \sim R_2$ means $G \ll K_1, K_2$.

   Then we need to translate G in this condition into totals and binding constants. This can be done for each dominance regime. So let us consider each dominance regime, namely $G_{\text{tot}} \sim G$, $G_{\text{tot}} \sim C_1$ and $G_{\text{tot}} \sim C_2$. For each dominance regime, re-express the behavior condition $G \ll K_1, K_2$ in terms of totals and binding constants.

   Now, some of these conditions involve the input variables $R_{1,\text{tot}}$ and $R_{2,\text{tot}}$, which can be disregarded, since these variables are varied freely across all positive reals, so there is always some region in the input space that satisfies conditions on them. So the validity condition for the behavior only needs to consider the conditions involving only the parameters, not the inputs. Show that the only conditions left, which is the validity condition for this behavior, is $G_{\text{tot}} \ll K_1, K_2$.

4. (Optional) What happens in the other behavior conditions? For example, what happens if $G_{\text{tot}} \gg K_1, K_2$? This corresponds to the limit where both R_1 and R_2 binds with G tightly. So you can consider the total number of tight binders $R_{\text{tot}} = R_{1,\text{tot}} + R_{2,\text{tot}}$, and G is a simple function of G_{tot} and R_{tot}, just like the tight binding limit of $G + R \rightleftharpoons C_{GR}$. Write down the specific expression of G as a function of $R_{1,\text{tot}}$ and $R_{2,\text{tot}}$. Note that there is a discontinuous or ultrasensitive transition between the two regimes in the tight binding limit.

   Then, what are the dominance regimes that this behavior includes? For each dominance regime included, look at their validity conditions, and see that indeed their overall behavior's validity condition is $G_{\text{tot}} \gg K_1, K_2$.

   What happens if $K_1 \ll G_{\text{tot}} \ll K_2$, so that G binds with R_1 tightly, but binds with R_2 weakly? We can again consider the R_2 overabundance limit, so $R_{2,\text{tot}} \sim R_2$. Then $G + R_1 \rightleftharpoons C_1$ can be treated as tight-binding limit, to have two dominance regimes. One regime is $G + C_1$ more than $R_{1,\text{tot}}$, in which case the dominance conditions are $G + C_1 \sim G$ and $R_{1,\text{tot}} \sim C_1$. In this regime, we can derive

$$G \sim G_{\text{tot}} \frac{1}{1 + \frac{R_{2,\text{tot}}}{K_2}}. \tag{4}$$

   The other regime is $G + C_1$ less than $R_{1,\text{tot}}$, in which case the dominance conditions are $G + C_1 \sim C_1$ and $R_{1,\text{tot}} \sim R_1$. In this regime, $G \sim (G + C_1)\frac{K_1}{R_{1,\text{tot}}}$. Based on this, you can derive

$$G \sim G_{\text{tot}} \frac{1}{\frac{R_{1,\text{tot}}}{K_1} + \frac{R_{2,\text{tot}}}{K_2}}. \tag{5}$$

   Derive the validity conditions for each of these two regimes. Compare this behavior with the previous one where both R_1 and R_2 are tight binding.

A relevant reference on such problems using traditional methods of analysis is Chapter 1 of the book [1].

1.2 Titration sponge (Optional)

Another example that is similar in binding network topology to the competitive binding example is the molecular titration sponge. Consider the following binding network:

$$G + R \rightleftharpoons C_{GR}, \quad R + T \rightleftharpoons C_{RT}. \tag{6}$$

Here G is a gene that can bind with a repressor R, while the repressor can also be sequestered by a titrator molecule T. The conserved quantities are $G_{\text{tot}} = G + C_{GR}$, $R_{\text{tot}} = R + C_{GR} + C_{RT}$, and $T_{\text{tot}} = T + C_{RT}$. The dissociation constants for the two binding reactions are denoted K_{GR} and K_{RT} respectively.

Since we are interested in the gene expression, and the repressor number tends to be dynamically adjusted while the titrator is used to change the overall behavior of the system, let us consider G as output, R_{tot} as input, and $T_{\text{tot}}, G_{\text{tot}}, K_{GR}, K_{RT}$ as parameters. Compared with competitive binding, we can map G to R_1, R to G, and T to R_2, and see that here we are effectively taking R_1 as output, and $R_{2,\text{tot}}$ as input, while $R_{1,\text{tot}}$ is also considered as a parameter.

The idea of this titration sponge behavior is the following. When R represses gene expression in the typical saturation behavior of a binding reaction $G + R \rightleftharpoons C_{GR}$, where we have

$$G \sim G_{\text{tot}} \frac{1}{1 + \frac{R_{\text{tot}}}{K_{GR}}}, \tag{7}$$

if we want gene expression to be sensitive to $R_{\text{tot}} = R + C_{GR}$, this requires $R_{\text{tot}} \gg K_1$. But we may want to elevate the threshold of repression, which is currently K_1, while keeping R's binding with G relatively tight. So we would like a different method to elevate the repression threshold, which can be done by titrating a number of repressors away by the titrator molecules T.

Let us investigate whether this can work and what are the conditions for it to hold.

1. Since we are interested in the repression of R on G, let us assume the R-saturation behavior for the binding reaction $R + G \rightleftharpoons C_{GR}$ between them. Show that this implies

$$G \sim G_{\text{tot}} \frac{1}{1 + \frac{R + C_{GR}}{K_{GR}}}. \tag{8}$$

2. Now we add in titrator. To be an effective titrator, we need the repressor to bind with the titrator first, so their binding reaction $R + T \rightleftharpoons C_{RT}$ should be tight-binding. Furthermore, since T serves as a titrator but not an inhibition on R, its number should be less than the number of repressors, which means $T_{\text{tot}} \ll R + C_{RT}$. Show that, under this condition, we have

$$C_{RT} \sim T_{\text{tot}}, \tag{9}$$

which in turn implies

$$R + C_{GR} \sim R_{\text{tot}} - T_{\text{tot}}. \tag{10}$$

These results together imply

$$G \sim G_{\text{tot}} \frac{1}{1 + \frac{R_{\text{tot}} - T_{\text{tot}}}{K_{GR}}}. \tag{11}$$

So we see that indeed this gives an effective elevation of the repression threshold. Compare the new threshold $R_{\text{tot}} - T_{\text{tot}} \gg K_{GR}$ with $R_{\text{tot}} \gg K_{GR}$.

3. What is the validity condition for this behavior? List the dominance regimes involved in this behavior, and derive their validity conditions.

4. Now, what if we assume that R binds with G tightly as well, but $K_{RG} \gg K_{RT}$ so that R binds with T even tighter? What would be the resulting function of G in terms of G_{tot}, R_{tot}, and T_{tot}? Note that there should be an ultrasensitive transition between the two regimes in this behavior. Compare this tight-binding titration sponge behavior to the previous saturation titration sponge behavior.

A reference on the titration effect in molecular interactions is this chapter of the online textbook on systems biology `https://biocircuits.github.io/chapters/12_molecular_titration.html`. Titration is also used in this 2016 work [2] to reduce noise and improve the original repressilator (which kick started synthetic biology in the year 2000) to persistently oscillate over hundreds of generations! Titration is also found to be the natural mechanism that controls replication initiation, to stably couple genome replication and cellular growth. In [3], it is shown that by titrating a replication initiator molecule with binding boxes on the genome, the cell can form a "progress bar" measuring how far replication has gone compared to growth, therefore coupling the two processes.

1.3 Tunability in two-layer regulations (Optional)

Previous two examples considered the same binding network topology, with two binding reactions connected "in parallel". In this example, let us consider connecting two binding reactions "in series". Consider the following binding network:

$$L + R \rightleftharpoons C_{LR}, \quad C_{LR} + P \rightleftharpoons C_{LRP}. \tag{12}$$

For a biological context, we can consider L as a ligand, binding with a receptor R to form an activated complex C_{LR}, which can then bind with a protein P to activate it. So the catalytic activity species is C_{LRP}. The conserved quantities are $L_{\text{tot}} = L + C_{LR} + C_{LRP}$, $R_{\text{tot}} = R + C_{LR} + C_{LRP}$, and $P_{\text{tot}} = P + C_{LRP}$. Let K_{LR} and K_{LRP} denote the dissociation constants of the two binding reactions.

The input is L_{tot}, the output is C_{LRP}, and the parameters are $R_{\text{tot}}, P_{\text{tot}}, K_{LR}, K_{LRP}$.

The idea of this behavior is that we want to regulate the activity of P by varying L_{tot}, and do this by first binding with a receptor, and then bind with the protein. We would like to see how does this additional step in the middle help with the diversity of the regulation, when compared with a direct activation of P by L.

1. Let us consider L is overabundant in its binding with R. As for the binding of C_{LR} with P, let us consider the regime where neither are saturated, so $C_{LR} + C_{LRP} \sim C_{LR}$, and $P_{\text{tot}} \sim P$. Show that under these assumptions, we have

$$C_{LRP} \sim \frac{P_{\text{tot}} R_{\text{tot}}}{K_{LRP}} \frac{L_{\text{tot}}/K_{LR}}{1 + L_{\text{tot}}/K_{LR}}. \tag{13}$$

   Compare this with the formula for a direct activation of P by L via binding $P + L \rightleftharpoons C_{PL}$ and have

$$C_{PL} \sim P_{\text{tot}} \frac{L_{\text{tot}}/K_{PL}}{1 + L_{\text{tot}}/K_{PL}} \tag{14}$$

   when L is overabundant.

   We see that adding the middle step of binding with R adds a knob $\frac{R_{\text{tot}}}{K_{LRP}}$ that adjusts the maximal activation. However, this can already be done by adjusting P_{tot}. Is there a situation where this behavior with R-binding step in the middle is more useful when compared with the direct activation?

2. Consider the behavior with the assumptions on the two binding reactions swapped. Namely, let us consider C_{LR}'s binding with P having C_{LR} overabundant, and consider L and R's binding with neither saturating. Show that in this case the behavior is

$$C_{LRP} \sim P_{\text{tot}} \frac{\frac{L_{\text{tot}} R_{\text{tot}}}{K_{LR} K_{LRP}}}{1 + \frac{L_{\text{tot}} R_{\text{tot}}}{K_{LR} K_{LRP}}}. \tag{15}$$

   Note that this can be expressed as

$$C_{LRP} \sim P_{\text{tot}} \frac{\frac{L_{\text{tot}}}{K_{\text{eff}}}}{1 + \frac{L_{\text{tot}}}{K_{\text{eff}}}}, \tag{16}$$

   where $K_{\text{eff}} = \frac{K_{LR} K_{LRP}}{R_{\text{tot}}}$ is the effective dissociation constant.

   Compare this with the direct activation behavior, we see that the R-binding step adds the R_{tot} knob that can tune the activation threshold. Given this, if we want to suppress the activity C_{LRP}, what can we do other than degrading L_{tot}?

3. List the dominance regimes involved in the three behaviors considered thus far: direct activation, overabundant L, and overabundant C_{LR}. Look at the dominance regimes' validity conditions, and derive the behaviors' validity conditions. Compare them and discuss each behavior is suited for what situations.

2 Adaptation and realizability

The idea about realizability is that in the world of synthetic biology, we design biomolecular circuits to achieve a certain function, such as adaptation. However, due to assumptions and approximations we make, some biocircuit designs may achieve the desired function in a wide range of conditions, therefore easily realized in experiments, while others may have a very limited range of conditions to achieve the desired function, therefore hard to realize in experiments. We can call this property the realizability of a given biocircuit. What complicates the problem even further is that we could get a wrong result on the realizability of a given circuit if our method of analysis is not careful. In this problem, we briefly explore this by looking into the realizability of a particular biocircuit design that achieves adaptation. This is based on the ongoing research work of Qinguo Liu.

In the landmark 2009 paper [4], the authors computationally explored a class of biocircuits that can achieve perfect adaptation. One such network topology they discovered is negative feedback. This network is illustrated in Figure 1.

This network has two species, A and B, that exist in both phosphorylated (A^*, B^*) and dephosphorylated (A, B) forms. The enzyme catalyzing the phosphorylation of A is the input species, I. Since the phosphorylated form of A is active, we can consider the total concentration of phosphorylated A, i.e. $A^*_{\text{tot}} = A^* + C_2 + C_3$, as output. So the goal is to have the concentration of output A^*_{tot} invariant to changes in the input $I_{\text{tot}} = I + C_1$, by feedback regulations from B's activities.

This is achieved by phosphorylations and dephosphorylation activities coupling A and B. The enzyme catalyzing the phosphorylation of B is a constant enzyme species, E. A and B's activities are coupled by A^* catalyzing B's phosphorylation, and B^* catalyzing A^*'s dephosphorylation. Note that, since all catalysis reactions are phosphorylations and dephosphorylations, the overall number of A and B molecules are conserved, i.e. $A_{\text{tot}} + A^*_{\text{tot}} = A + A^* + C_1 + C_2 + C_3$ and $B_{\text{tot}} + B^*_{\text{tot}} = B + B^* + C_2 + C_3 + C_4$ do not change with time. So we can describe the dynamics of the system by two variables, which we can choose to be $A^*_{\text{tot}} = A^* + C_2 + C_3$ and

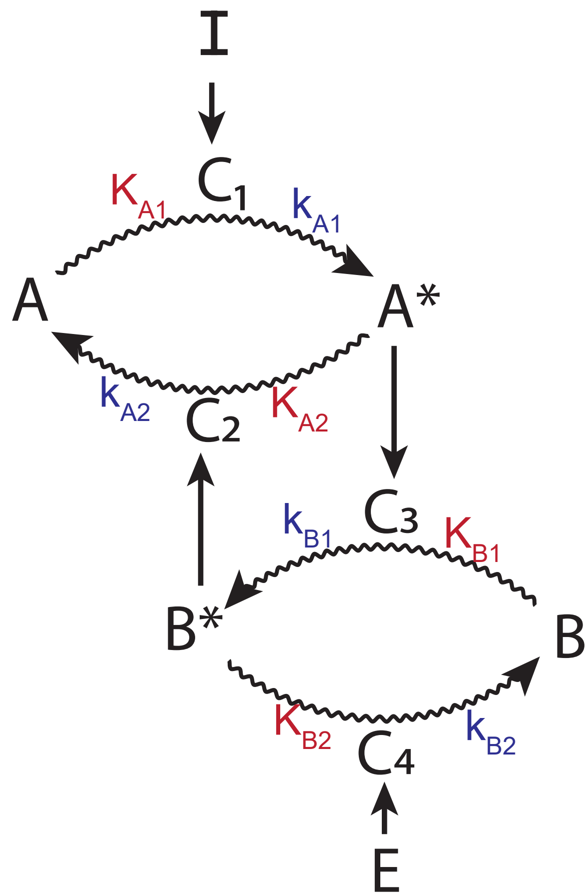

Figure 1 Illustration of the negative feedback loop topology that achieves perfect adaptation, as depicted in [4]. The blue small k parameters are catalysis rate constants, and the red K parameters are dissociation constants for the binding reactions. Credit to Qinguo Liu.

$B^*_{\mathrm{tot}} = B^* + C_2 + C_4$. This yields the following dynamical system,

$$\begin{aligned} \frac{d}{dt}A^*_{\mathrm{tot}} &= k_{A1}C_1 - k_{A2}C_2, \\ \frac{d}{dt}B^*_{\mathrm{tot}} &= k_{B1}C_3 - k_{B2}C_4, \end{aligned} \tag{17}$$

where the complexes are formed by the following binding network,

$$I + A \rightleftharpoons C_1, \;\; A^* + B^* \rightleftharpoons C_2, \;\; B + A^* \rightleftharpoons C_3, \quad B^* + E \rightleftharpoons C_4. \tag{18}$$

2.1 An analysis of adaptation

The goal of the system is to have the steady state concentration of A^*_{tot} be invariant to changes in the input I_{tot}. Let us first see how this can be true, following arguments adapted from [4].

1. The key idea is to let B^*_{tot} serve as an integral variable of A. So let us focus on the two binding reactions forming C_3 and C_4, which determines the dynamics of B^*_{tot}.

Let us assume that there exists one regime such that $C_3 \sim A^*_{\text{tot}}$ and $C_4 \sim E_{\text{tot}}$. Write down the equation for the dynamics of B^*_{tot} under this regime, and argue that if steady state is achieved, the output A^*_{tot} is determined solely by parameters independent of the input $I_{\text{tot}} = I + C_1$.

2.2 A traditional analysis for validity conditions that fails... (Optional)

How to derive conditions that achieve the above desired integral feedback for perfect adaptation? Traditionally in systems and synthetic biology, we assume enzymatic reactions have saturation behaviors of the Michaelis-Menten form, which can dramatically simplify the regulatory behavior of a binding network. Then, further simplifying assumptions can be made to see what conditions can yield the desired function. This is also the approach taken in [4].

Let us follow this approach, but pay close attention to the conditions needed to assume enzymatic reaction rates take Michaelis-Menten form. The assumptions made when writing down Michaelis-Menten functions are often ignored in research works in this field, but it could result in catastrophic failures, as we will see at the end of this problem.

1. First, let us assume every catalysis reaction has substrates saturating, so the catalysis rates take Michaelis-Menten form. This means, for the enzyme-substrate binding reaction $E + S \rightleftharpoons C$ with dissociation constant K, we assume $E_{\text{tot}} \ll K$ so that $C \sim E_{\text{tot}} \frac{S_{\text{tot}}}{S_{\text{tot}}+K}$.

   (A reminder on our analysis of dominance regimes. There are exactly three dominance regimes in one binding reaction $E + S \rightleftharpoons C$, with dissociation constant K. One is $E + C \sim E$, $S + C \sim S$, with $C \sim \frac{(E+C)(S+C)}{K}$. Another is $E + C \sim C$, $S + C \sim S$, with $C \sim E + C$. The last is $E + C \sim E$, $S + C \sim C$, with $C \sim S + C$. The fact that there are exactly three regimes in a binding reaction is very powerful and is used throughout for the analysis of bioregulations.)

   By our design, B^*_{tot} should be the integral variable of A^*_{tot}. Therefore, C_2, C_3 and C_4, which involve B and B^*, play the central role in the feedback regulation of B on A. So we focus on the binding reactions of C_2, C_3 and C_4. Let us walk through the Michaelis-Menten assumption for C_2.

   The binding reaction for C_2 is $A^* + B^* \rightleftharpoons C_2$, with dissociation constant K_{A2}, and B^* is the enzyme and A^* is the substrate. For Michaelis-Menten behavior, we would like to write

   $$C_2 \sim B^*_{\text{tot}} \frac{A^*_{\text{tot}}}{A^*_{\text{tot}} + K_{A2}}. \tag{19}$$

   To do so, we first assume $B^* + C_2 \ll K_{A2}$ so that the A^*-saturation behavior holds. So we have

   $$C_2 \sim (B^* + C_2) \frac{(A^* + C_2)}{(A^* + C_2) + K_{A2}}.$$

   To replace $B^* + C_2$ with $B^*_{\text{tot}} = B^* + C_2 + C_4$, we need to assume $C_4 \ll B^* + C_2$. To replace $A^* + C_2$ with $A^*_{\text{tot}} = A^* + C_2 + C_3$, we need to assume $C_3 \ll A^* + C_2$. So the conditions for this behavior is

   $$C_4 \ll B^* + C_2 \ll K_{A2}, \quad C_3 \ll A^* + C_2. \tag{20}$$

   Perform a similar analysis for the binding reactions of C_3 and C_4 to obtain the following desired behaviors and corresponding conditions.

   $$C_3 \sim A^*_{\text{tot}} \frac{B_{\text{tot}}}{B_{\text{tot}} + K_{B1}}, \qquad C_2 \ll A^* + C_3 \ll K_{B1}. \tag{21}$$

$$C_4 \sim E_{\text{tot}} \frac{B^*_{\text{tot}}}{B^*_{\text{tot}} + K_{B2}}, \qquad E_{\text{tot}} \ll K_{B2}, \;\; C_2 \ll B^* + C_4. \tag{22}$$

2. Combine the conditions for C_2 and C_3 to obtain that the overall condition is

$$C_2, C_4 \ll B^* \ll K_{A2}, \;\; C_2, C_3 \ll A^* \ll K_{B1}, \;\; E_{\text{tot}} \ll K_{B2}. \tag{23}$$

   In particular, notice that the condition $C_2, C_4 \ll B^*$ implies $B^*_{\text{tot}} = B^* + C_2 + C_4 \sim B^*$. Similarly, we have $A^*_{\text{tot}} = A^* + C_2 + C_3 \sim A^*$.

   Based on this, we have already specified the dominance regime for the binding reaction $A^* + B^* \rightleftharpoons C_2$. Argue that $C_2 \sim \frac{A^*_{\text{tot}} B^*_{\text{tot}}}{K_{A2}}$. Then show that $C_2 \ll B^*$ implies $A^*_{\text{tot}} \ll K_{A2}$.

   Similarly, we have also specified the dominance regime for the binding reaction $A^* + B \rightleftharpoons C_3$, since C_3 does not dominate $B + C_3$ or $A^* + C_3$. Argue that $C_3 \sim \frac{A^*_{\text{tot}} B_{\text{tot}}}{K_{B1}}$. Then show that $C_3 \ll A^*$ implies $B_{\text{tot}} \ll K_{B1}$.

   The following summarizes the conditions we have obtained thus far in terms of totals:

$$B_{\text{tot}} \ll K_{B1}, B^*_{\text{tot}} \ll K_{A2}, \;\; A^*_{\text{tot}} \ll K_{B1}, K_{A2}, \;\; E_{\text{tot}} \ll K_{B2}. \tag{24}$$

3. Now, recall that, under the desired saturation (Michaelis-Menten) behaviors, we have the following dynamics of B^*_{tot}:

$$\frac{d}{dt} B^*_{\text{tot}} = k_{B1} C_3 - k_{B2} C_4 \approx k_{B1} A^*_{\text{tot}} \frac{B_{\text{tot}}}{B_{\text{tot}} + K_{B1}} - k_{B2} E_{\text{tot}} \frac{B^*_{\text{tot}}}{B^*_{\text{tot}} + K_{B2}}. \tag{25}$$

   To achieve the desired result that B^*_{tot} serves as an integral variable of A^*_{tot}, we need $B_{\text{tot}} \gg K_{B1}$ and $B^*_{\text{tot}} \gg K_{B2}$.

   Does this contradict the conditions we have assumed previously? What does this imply?

2.3 Holistic regimes come to rescue

In our previous analysis, we followed the traditional approach where we first make the assumption that all enzymatic reactions follow Michaelis-Menten kinetics, and then find regimes where our desired behavior is achieved. This eventually resulted in contradictory conditions such that seemingly our desired behavior can NEVER be achieved!

However, the fault in this case actually rests on our method of analysis. Making the Michaelis-Menten assumption in the first step already ruled out several regimes, which includes regimes that achieves perfect adaptation.

We can "rescue" our previous analysis by getting rid of Michaelis-Menten assumptions altogether, and directly search for the regimes that achieve our desired behavior. We do this below.

1. Based on the idea about this adaptation behavior, we want a regime where B^*_{tot} serves as the integral variable of A^*_{tot}. Therefore, we want $C_3 \sim A^*_{\text{tot}}$ and $C_4 \sim E_{\text{tot}}$. We can start with this to derive the dominance regimes that may achieve our desired function.

Let us list the total variables so we have an idea of all possible regimes.

$$\begin{aligned}
I_{\text{tot}} &= I + C_1, \\
A_{\text{tot}} &= A + C_1, \\
A^*_{\text{tot}} &= A^* + C_2 + C_3, \\
B_{\text{tot}} &= B + C_3, \\
B^*_{\text{tot}} &= B^* + C_2 + C_4, \\
E_{\text{tot}} &= E + C_4.
\end{aligned} \tag{26}$$

Let us start with the binding reactions of C_3 and C_4. What is the desired regime for $B + A^* \rightleftharpoons C_3$? And what is the desired regime for $B^* + E \rightleftharpoons C_4$?

Show that, under these regimes, the total variables satisfy

$$\begin{aligned}
I_{\text{tot}} &= I + C_1, \\
A_{\text{tot}} &= A + C_1, \\
A^*_{\text{tot}} &\sim C_3, \\
B_{\text{tot}} &\sim B, \\
B^*_{\text{tot}} &\sim B^* + C_2, \\
E_{\text{tot}} &\sim C_4.
\end{aligned} \tag{27}$$

2. So we can choose one dominance regime to check whether the desired integral feedback property holds. Let us consider the regime where $(I_{\text{tot}}, A_{\text{tot}}, A^*_{\text{tot}}, B_{\text{tot}}, B^*_{\text{tot}}, E_{\text{tot}}) \sim (I, A, C_3, B, C_2, C_4)$.

   Show that in this case the system dynamics is

$$\begin{aligned}
\frac{d}{dt}A^*_{\text{tot}} &= k_{A1}\frac{I_{\text{tot}}A_{\text{tot}}}{K_{A1}} - k_{A2}B^*_{\text{tot}}, \\
\frac{d}{dt}B^*_{\text{tot}} &= k_{B1}A^*_{\text{tot}} - k_{B2}E_{\text{tot}}.
\end{aligned} \tag{28}$$

   Use the condition that $A_0 = A^*_{\text{tot}} + A_{\text{tot}}$ is conserved to rewrite the system as the following:

$$\frac{d}{dt}\begin{bmatrix} A^*_{\text{tot}} \\ B^*_{\text{tot}} \end{bmatrix} = \begin{bmatrix} -k_{A1}\frac{I_{\text{tot}}}{K_{A1}} & -k_{A2} \\ k_{B1} & 0 \end{bmatrix} \begin{bmatrix} A^*_{\text{tot}} \\ B^*_{\text{tot}} \end{bmatrix} + \begin{bmatrix} k_{A1}\frac{I_{\text{tot}}}{K_{A1}}A_0 \\ -k_{B2}E_{\text{tot}} \end{bmatrix}. \tag{29}$$

   Show that this system has a unique fixed point, which satisfies $A^*_{\text{tot}} = \frac{k_{B2}}{k_{B1}}E_{\text{tot}}$, and this fixed point is stable. So this system achieves perfect adaptation.

3. (Optional.) For the dominance regime selected above, derive its validity condition, which should be in terms of totals and dissociation constants.

   This condition corresponds to one region in parameter space that the perfect adaptation behavior holds. Compare this dominance regime and the validity conditions with our analysis in the previous subproblem where we first made Michaelis-Menten assumptions. Is this regime eliminated when making Michaelis-Menten assumption? If we only make Michaelis-Menten assumption on some of the binding reactions, which subset of binding reactions can be assumed as Michaelis-Menten and still keep this regime included?

3 Glycolytic oscillations and robustness-efficiency tradeoffs of adaptations

We can use perspectives of control theory to better understand the design principles of natural biological systems. We do so by conceptualizing a plant-controller split, and then by analyzing what is special about the natural controller among the space of all possible controller designs, we can get a glimpse of why natural biological systems evolved into this form. Of course, if the natural controller does not seem special and indeed inferior to some of the controller designs, then instead of lamenting on evolution is not optimal or life is not designed, we can always put our engineer hat on and construct synthetic biological systems that surpass the performance of natural ones!

To illustrate this perspective of using control theory to understand biological design principles, let us investigate the control of the glycolysis pathway. The glycolysis pathway consumes glucose and produces protons, ATPs, and metabolic intermediates that are part of the central metabolism and connects to many other metabolic pathways. It is so central that it is present in essentially every biological organism. Therefore, it is reasonable to consider the hypothesis that the regulation of the glycolysis pathway is optimized in some sense.

Since the early 1960s, it has been observed that when cells are starved, and glucose is suddenly added, oscillations in the concentrations of the glycolysis pathway's intermediates (namely NADH) at minutes timescale has been observed [5]. People hypothesized that this must serve some functional purpose. But intuitively, we would imagine that the main objective of controlling the glycolysis pathway is to stably supply ATP, protons, and other intermediates, not oscillating them. Also, such suspicion is further supported by the fact that such oscillations are only observed under extreme and unrealistic scenarios. Therefore, could it be that glycolytic oscillations do not serve any functional purposes? But then why cells do not get rid of it? In the 2011 work [6], through control theory analysis, it is shown that glycolytic oscillations is simply an inevitable side effect of the controller design that adapts to changing demands on ATPs. In particular, there is a fundamental tradeoff between robustness and efficiency that cannot be broken for all possible controller designs. If a controller adapts to changing ATP demands efficiently, then it inevitably tends to oscillate and is fragile.

We follow some parts of the analysis in [6] in this problem.

3.1 Steady state analysis of the tradeoff in a simplified glycolysis model

While the actual glycolysis pathway consists of many steps, given our goal of analyzing glycolytic oscillations, we can focus on ATP production with overabundant glucose supply. Therefore, we can simplify the glycolysis pathway into just two steps. The first step is one unit of ATP is consumed to activate glucose and produce one unit of the intermediate molecule. The second step is the consumption of one intermediate molecule to produce two units of ATP.

$$\mathrm{ATP} \rightsquigarrow \mathrm{Intermediate}, \qquad \mathrm{Intermediate} \rightsquigarrow 2\mathrm{ATP}. \tag{30}$$

Since ATP is supplied by glycolysis to other metabolic pathways, we also need to consider the consumption of ATP. So we have another reaction

$$\mathrm{ATP} \rightsquigarrow \emptyset. \tag{31}$$

1. Denote $x_1 = x_{\text{int}}$ as the concentration of intermediate, and $x_2 = x_{\text{ATP}}$ as the concentration of intermediate. Then denote $v_1 = v_{\text{PFK}}$ and $v_2 = v_{\text{PK}}$ as the reaction fluxes for the two catalysis reactions, since one step among the ATP-consuming reactions is known to be catalyzed by the PFK enzyme, and one step among the ATP-producing reactions is known to be catalyzed by the PK enzyme. The third reaction, the consumption of ATP, can be viewed as disturbance flux w, since the consumption of ATP is determined by other parts of the cell and the external environment, which can increase when experiencing hardship such as heat shock. Write the system in the following form:

$$\frac{d}{dt}x = Sv + S^w w, \tag{32}$$

where S is the reaction stoichiometry matrix of the two catalysis reactions, and v is the two-dimensional vector of the two reaction's fluxes, S^w is the stoichiometry matrix of the third reaction, and w is the scalar flux of the third reaction. What are the matrices S and S^w?

2. To make this system have internal metabolic dynamics, we need to know how the fluxes v are regulated by the metabolites x. In this particular case, from biochemical experiments, we know that the PFK and PK enzymes are allosterically activated by AMP, therefore effectively inhibited by ATP. So we can assume they take the following form:
$$v_1 = v_{\text{PFK}} = \frac{2x_{\text{ATP}}^a}{1 + x_{\text{ATP}}^{2h}}, \qquad v_2 = v_{\text{PK}} = \frac{2kx_{\text{int}}}{1 + x_{\text{ATP}}^{2g}}, \tag{33}$$
where a is the sensitivity of v_{PFK} to x_{ATP}, with a typical value of $a = 1$, h is the strength of inhibition of this flux by ATP, with a typical value from 1 to 4, k is the rate constant for v_{PK}, and g is the strength of inhibition of the PK flux by ATP, with a typical value between 0 and 1. There are "2"s here and there, for the purpose of normalization and simplification.

   Write down the full control system given how the fluxes are regulated by the metabolites. Assume the reference value of the disturbance flux w is $w^* = 1$. Then, with w kept constant, this control system becomes an autonomous dynamical system, amenable to stability analysis.

   Show that the steady state equations are $v_1(x^*) - v_2(x^*) = 0$ and $2v_2(x^*) - v_1(x^*) - w^* = 0$, where x^* is the steady state concentration of x. Solve this to get that one fixed point is
$$x^* = (x_1^*, x_2^*) = (x_{\text{int}}^*, x_{\text{ATP}}^*) = (1/k, 1). \tag{34}$$
   This fixed point is unique if $a = 2h$, for example.

   Then, denote $\Delta x = x - x^*$ as the deviation from the fixed point, and similarly denote $\Delta w = w - w^*$, and linearize the system around this fixed point. Show that the linearized reaction fluxes are
$$\Delta v_{\text{PFK}} = (a - h)\Delta x_{\text{ATP}}, \qquad \Delta v_{\text{PK}} = k\Delta x_{\text{int}} - g\Delta x_{\text{ATP}}. \tag{35}$$
   And the linearized system is the following:
$$\frac{d}{dt}\begin{bmatrix} \Delta x_{\text{int}} \\ \Delta x_{\text{ATP}} \end{bmatrix} = \begin{bmatrix} -k & a+g-h \\ 2k & -a-2g+h \end{bmatrix}\begin{bmatrix} \Delta x_{\text{int}} \\ \Delta x_{\text{ATP}} \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix}\Delta w. \tag{36}$$

3. Analyze the stability of the fixed point using the Routh-Hurwitz criterion. (Note that $k > 0$.) Show that the stability condition can be expressed as
$$0 < h - a < k + 2g. \tag{37}$$

4. The output of the system is ATP concentration, since we care about maintaining ATP at a reference level. Therefore the steady state error in response to a given perturbation can be expressed as $\left|\frac{\Delta x_{\text{ATP}}}{\Delta w}\right|$.

   Show that the steady state error satisfies
$$\left|\frac{\Delta x_{\text{ATP}}}{\Delta w}\right| = \left|\frac{1}{h-a}\right|. \tag{38}$$

   Observe that there is a tradeoff between the steady state error and the margin of stability. If we want to adapt to changing demands, i.e. changing Δw, we need the steady state error to be small. This means we want a large $h - a$, which requires a large $k + 2g$ to keep the system stable. Since k corresponds to the number of enzymes, a large $k + 2g$ means a larger metabolic overhead to produce a large number of enzymes. From this, we see that to adapt to a changing environment, there is a tradeoff between robustness, i.e. margin of stability, and efficiency, i.e. metabolic overhead to produce enzymes.

This also explains that sustained oscillations, which happens when $k + 2g$ goes below $h - a$ to make the system unstable, happens under starvation causing the number of enzymes to be low so that k is small. This shows that glycolytic oscillation may not have a functional role in itself, but a side effect of having an $h - a$ such that the system adapts to changing demands on ATP with a small steady state error, which only becomes apparent under conditions such as starvation causing k to be too small.

3.2 Simulations reveal dynamic tradeoffs for varying control parameters (Optional)

Our previous analysis reveals a simple tradeoff between robustness and efficiency when the system adapts to changing environments at steady state. Through simulations, we can get a sense of how the tradeoff on system performance also holds dynamically.

Take the full (nonlinear) system, and run numerical simulations of the system trajectory. Try $g = 0$, and increase h from 1 to 4. What do you observe? Then choose an h with small enough steady state error, and increase k, what do you observe? Note that large oscillations means the system is fragile and sensitive to disturbances.

Try $g = 1$ and do the same increase on h. How does the trajectories compare with the $g = 0$ case? What would you imply from this?

References

[1] J. Keener and J. Sneyd, *Mathematical Physiology I.* Springer-Verlag New York, 2009.

[2] L. Potvin-Trottier, N. D. Lord, G. Vinnicombe, and J. Paulsson, "Synchronous long-term oscillations in a synthetic gene circuit," *Nature*, vol. 538, 2016.

[3] H. Fu*, F. Xiao*, and S. Jun, "Bacterial replication initiation as precision control by protein counting," *PRX Life, **highlighted with Viewpoint***, vol. 1, p. 013011, Aug 2023. [Online]. Available: https://link.aps.org/doi/10.1103/PRXLife.1.013011

[4] W. Ma, A. Trusina, H. El-Samad, W. A. Lim, and C. Tang, "Defining network topologies that can achieve biochemical adaptation," *Cell*, vol. 138, no. 4, pp. 760–773, 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0092867409007120

[5] S. Danø, P. G. Sørensen, and F. Hynne, "Sustained oscillations in living cells," *Nature*, vol. 402, no. 6759, pp. 320–322, Nov 1999. [Online]. Available: https://doi.org/10.1038/46329

[6] F. A. Chandra, G. Buzi, and J. C. Doyle, "Glycolytic oscillations and limits on robust efficiency," *Science*, vol. 333, no. 6039, pp. 187–192, 2011.