1. **Binding kinetics, detailed balance and cooperativity**

*D* is a DNA segment containing two sites, *A* and *B*, to which the protein *X* can bind in any order.

![Diagram of binding kinetics](image)

(a) In the figure, the *K* *i* represent association constants in inverse concentration units. According to detailed balance, the individual binding reactions are in equilibrium. Find two different expressions for the equilibrium ratio $[DX_AX_B]/[D][X]^2$, corresponding to the two possible binding paths. Can all four association constants be independently specified?

(b) The total DNA concentration is $[D_{tot}] = [D] + [DX_A] + [DX_B] + [DX_AX_B]$. If $K_2 >> K_1$ and $K_4 >> K_3$, show that the singly-bound forms $[DX_A]$ and $[DX_B]$ form a negligible fraction of this total. This corresponds to a cooperative system in which the second binding is highly encouraged once the first has already occurred.

(c) Assuming $[D_{tot}] \approx [D] + [DX_AX_B]$, find the bound fraction $f = [DX_AX_B]/[D_{tot}]$ in terms of $[X]$ and the association constants.

2. **Dilution of proteins due to cell growth**

A single bacterial cell at time $t = 0$ has volume $V_0$. After a time interval $T_D$, the doubling time, the cell grows and divides into two cells, each of volume $V_0$; after another interval $T_D$, there are four cells, and so on.

(a) Show that the combined volume of cells at time $t$ may be written as $V(t) = V_0 e^{\gamma t}$. Find $\gamma$ in terms of $T_D$.

(b) The protein *X* is created at some rate $k(t)$, so the total number of molecules of *X* satisfies $\frac{dn_X}{dt} = k(t)$. Show that the concentration $[X] = n_X/V$ satisfies

$$\frac{d}{dt}[X] = \frac{k(t)}{V} - \gamma[X]$$

Discuss the origin of the decay term.
3. Positive feedback and bistability

Suppose the protein $X$ is a transcriptional activator, and $D$ a promoter which is activated by the binding of $X$. If the downstream gene happens to code for $X$ itself, the resulting positive feedback can lead to bistability.

\[
\text{DX}_A X_B
\]

\[
D
\]

(8) a. Let $v_I$ be the rate of expression from bound DNA ($DX_A X_B$), and $v_0 < v_I$ the rate of expression from free DNA ($D$); note that these rates of expression are assumed to have units of $T^{-1} \text{Conc}^{-1}$. Use the results of Problems 1 and 2 to show that the time-evolution of the concentration $x \equiv [X]$ may be written in the following form:

\[
\frac{dx}{dt} = \frac{v_0 + v_I K_1 K_2 x^2}{1 + K_1 K_2 x^2} - \gamma \cdot x
\]

(10) b. Steady state solutions occur at those values of $x$ for which the rates of generation ($f(x)$) and degradation ($g(x)$) are equal. Set $\gamma = 1$, $K_1 K_2 = 1$, and use the accompanying MATLAB file ps1_1.m to explore the intersections of $f(x)$ and $g(x)$ as the parameters $v_0$ and $v_I$ are varied. The figures (i) through (v) show schematically the types of behaviors that can occur. Plot an example of types (i), (iii) and (v) indicating the parameter values that generated them. Label those values of $x$ for which $f(x) > g(x)$ with a rightward arrow, and those for which $f(x) < g(x)$ with a leftward arrow. Which solutions are stable?

(8) c. The boundary between bistability and monostability is given by parameters for which the system has precisely two fixed points (types (ii) and (iv)). Setting $\gamma = 1$ and $K_1 K_2 = 1$, rewrite the condition $f(x) = g(x)$ as a cubic equation of the form $x^3 + c_2 x^2 + c_1 x + c_0 = 0$. Any cubic can always be factorized as $(x - a_1)(x - a_2)(x - a_3)$. What are the conditions on the roots that would lead to type (ii) or type (iv) behavior? Apply this condition and compare coefficients to obtain parametrized equations for $v_I$ and $v_0$. Show on a graph of $v_I$ vs. $v_0$ the region over which the system is bistable.

(4) d. Find examples of a real biological systems (other than the Lambda phage) where a protein activates its own production and one where multiple binding sites act cooperatively. Do those systems exhibit switching behavior?
4. Bistability in the *S. cerevisiae* GAL system, a toy model

In this problem we will explore a model of the GAL system in *S. cerevisiae* (budding yeast) in which bistability does not arise because of cooperative binding. The GAL system is a network that is activated by the sugar galactose (G) and consists of a chain of interactions between different proteins and protein complexes which are depicted in the adjacent figure.

You can assume that the only slow dynamical variable is the concentration of Gal3, the protein encoded by the *GAL3* gene. In the cytoplasm, galactose, which in this problem is assumed to be held at a fixed concentration, can bind to Gal3 forming the G-Gal3 complex. This compound, in turn, interacts with the Gal80c proteins present in the cytoplasm leading to the formation of a new complex G-Gal3-Gal80c. This effectively reduces the amount of Gal80 proteins that can be imported into the cell nucleus (Gal80n). Once in there Gal80n proteins can bind to the transcriptional activator Gal4 (which is constitutively bound to some DNA transcription sites), converting active transcription sites (ATS) into sites where transcription cannot be initiated (we will denote these sites by Gal80n-ITS to indicate that the binding of Gal80n is the cause of site inactivation). Active transcription sites (ATS) lead, on a slow timescale, to the synthesis of Gal3 proteins in the cytoplasm. The accompanying figure is a scheme of the process.

5. a. Do Gal3 proteins enhance or inhibit their own synthesis?

5. b. Assume that the system is being observed at a timescale slow compared to transcription and translation. Additionally assume that at this time scale all the other reactions rapidly approach equilibrium. Write down the 4 equations that represent the equilibrium of all 4 rapid reactions (\(K_1, K_2\) and \(K_4\) are dissociation constants, measured in units of concentration and \(k_3\) is defined as \(k_3 = k_{[\text{Gal80}_n \rightarrow \text{Gal80}_c]} / k_{[\text{Gal80}_c \rightarrow \text{Gal80}_n]}\) where \(k_{[\text{Gal80}_n \rightarrow \text{Gal80}_c]}\) and \(k_{[\text{Gal80}_c \rightarrow \text{Gal80}_n]}\) are the rate constants for Gal80 nuclear export and import respectively).

5. c. There are two concentrations that are conserved in the system. Write the equations that reflect this fact. At short timescales when protein synthesis can be neglected, there is an extra conserved quantity: the total concentration of all the forms in which the Gal3 protein participates; express this as an extra equation.
d. The total concentration of all the forms of Gal3 is called \([\text{Gal3}]_{\text{tot}}\). By using the equations deduced in b and c prove that the fraction \(f\) of active transcription sites satisfies the equation

\[
1 + k_3 \left( 1 + \frac{[G][\text{Gal3}]_{\text{tot}}}{K_1 K_2} \right) + \frac{[A]}{K_4} f = \frac{[B]}{K_4} \frac{f}{1 - f}
\]

What do the constants \(A\) and \(B\) represent?

It is quite difficult to solve this equation analytically but we can use MATLAB to solve it numerically. In what follows, assume \(K_1 = 1\%\) w/v (weight/volume); \(K_2 = 0.05\) nM; \(k_3 = 1\); \(K_4 = 0.05\) nM; \([A] = 50\) nM.

e. Assuming that the synthesis of Gal3 occurs at a rate proportional to the fraction of active transcription sites we can model the dynamics of the system at a slow timescale using the equation

\[
\frac{d[\text{Gal3}]_{\text{tot}}}{dt} = \kappa \cdot f([\text{Gal3}]_{\text{tot}}) - \gamma \cdot [\text{Gal3}]_{\text{tot}}
\]

where \(f([\text{Gal3}]_{\text{tot}})\) is the solution of the expression deduced in d for a given value of \([G]\) and \([B]\). Assuming \(\kappa / \gamma = 25000\) nM and using the accompanying MATLAB code \texttt{ps1}\_2.m (and \texttt{ps1}\_2\_fLHS.m) explore the bistable properties of the system in the range \([G]\in[0, 1]\%\) w/v and \([B]\in[0, 10000]\) nM. In particular, say whether the system is bistable or monostable in these configurations:

I. \([G] = 0.1\%\) w/v; \([B] = 25\) nM.
II. \([G] = 0.1\%\) w/v; \([B] = 250\) nM.
III. \([G] = 0.1\%\) w/v; \([B] = 2500\) nM.

f. In this system, what network element is the most relevant for establishing bistability?

(0) \(g\). CHALLENGE: Compute the boundaries of the bistability region in the \(([G], [B])\) space.