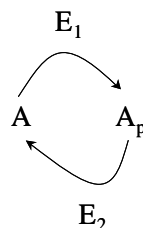


Problem 1:

Multiple intracellular processes are controlled by reversible phosphorylation of proteins. Kinases and phosphatases add and remove phosphate groups to specific aminoacids on substrate proteins. The phosphorylation and dephosphorylation reactions can be considered irreversible.

Consider a simple enzymatic network, where a substrate is interconverted between the phosphorylated and unphosphorylated states by two enzymes (kinase and phosphatases):



The total amount of the substrate is A_{total} . The phosphorylation and dephosphorylation follow Michaelis-Menten kinetics. E.g., the rate of the forward reaction is given by:

$$R_{phosphorylation} = \frac{V_1 A}{K_1 + A} \quad R_{dephosphorylation} = \frac{V_2 A_p}{K_2 + A_p}$$

Consider the case when the $K_1 \approx K_2 \equiv K$, and analyze the dependence of the phosphorylated fraction of substrate (A_p / A_{total}) on the ratio of the maximal rates of the forward and reverse reactions V_1 / V_2 . Plot the predicted dependences of A_p / A_{total} on V_1 / V_2 for several values of K . Interpret the observed behavior.

You can use the following papers in solving this problem:

1) [Goldbeter A, Koshland DE Jr.](#), An amplified sensitivity arising from covalent modification in biological systems. Proc Natl Acad Sci U S A. 1981 Nov;78(11):6840-4.

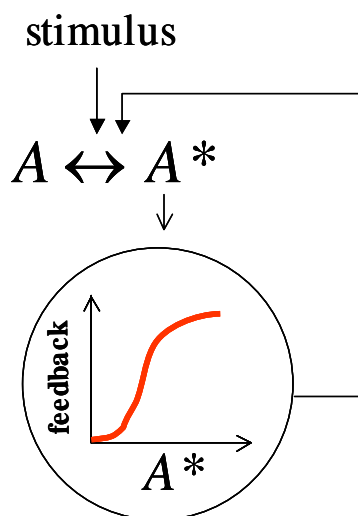
2) [Goldbeter A, Koshland DE Jr.](#), Ultrasensitivity in biochemical systems controlled by covalent modification. Interplay between zero-order and multistep effects. J Biol Chem. 1984 Dec 10;259(23):14441-7.

Problem 2:

Read the following papers on biochemical switches:

1. Ferrell JE, Xiong W. Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible. *Chaos*. 2001 Mar;11(1):227-236.
2. [Ferrell JE Jr.](#) Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr Opin Cell Biol*. 2002 Apr;14(2):140-8.

Consider the following network with a positive feedback loop:



The positive feedback can be described by a Hill function with Hill coefficient n . The dynamic balance equation for the active form of the protein can be written as:

$$\frac{dA^*}{dt} = k_1 S(A_T - A^*) + k_2 (A_T - A^*) \frac{A^{*n}}{\Gamma + A^{*n}} - k_{-1} A^*$$

- Explain different terms in this equation.
- Determine the key dimensionless groups that characterize the steady state behavior in this problem.
- Derive the condition for steady state multiplicity. (This condition should be in the form of an algebraic equation linking the dimensionless groups from the previous step).
- Develop an algorithm to compute the steady states in this problem as a function of the stimulus (exogenous input) to the network.

Problem 3:

- A. Download the cancer vs. normal gene expression data set for lung cancer (see course web page). This data set contains 15 tumor samples and 8 normal samples (one sample per column). Rows correspond to genes. Implement a suitable statistical test to identify genes that are overexpressed in tumors as compared to normal samples (attach your code to your submission, can be in any programming language or matlab or R). What are the top 25 genes you found? Is there any way you can do a "sanity check" to see that you found genes that are actually differentially expressed (describe how you would do it)? (submit: list of genes, your code, and description a "sanity check" test)
- B. Describe (don't implement) an quantitative evaluation scheme for your statistical test (either on biological or synthetic data or in some other way - up to you). Do not go into minute details, tell us what your gold standard is in this evaluation, why you chose to evaluate this way, what type of results you'd report. Does your evaluation have any drawbacks - what are they? (submit: description of your evaluation scheme)
- C. Does your statistical test in (A) depend on any assumptions? What are they? Discuss how you could design a different test that does not make those assumptions. What are the advantages and drawbacks of your new test?

Some possible references (you don't have to read all of these or implement algorithms that are as complicated as the ones suggested in these papers, but if you are lost, these may give you some ideas):

Thomas JG, Olson JM, Tapscott SJ, Zhao LP. An efficient and robust statistical modeling approach to discover differentially expressed genes using genomic expression profiles. *Genome Res.* 2001 Jul;11(7):1227-36.

Ideker T, Thorsson V, Siegel AF, Hood LE. Testing for differentially-expressed genes by maximum-likelihood analysis of microarray data. *J Comput Biol.* 2000;7(6):805-17.

Park T, Yi SG, Lee S, Lee SY, Yoo DH, Ahn JI, Lee YS. Statistical tests for identifying differentially expressed genes in time-course microarray experiments. *Bioinformatics.* 2003 Apr 12;19(6):694-703.