20.320 Problem Set #6

Due on Friday November 4th, 2011 at 11:59am. No extensions will be granted.

General Instructions:

1. You are expected to state all of your assumptions, and provide step-by-step solutions to the numerical problems. Unless indicated otherwise, the computational problems may be solved using Python/MATLAB or hand-solved showing all calculations. The results of any calculations must be printed and attached to the solutions, and the corresponding code should be submitted on Ô[˘ ¦•^Á ^à•æ^. For ease of grading (and in order to receive partial credit), your code must be well organized and thoroughly commented with meaningful variable names.
2. You will need to submit the solutions to **each problem to a separate mail box**, so please prepare your answers appropriately. Staple the pages for each question separately and make sure **your name** appears on each set of pages. (The problems will be sent to different graders, which should allow us to get graded problem sets back to you more quickly).
3. Submit your completed problem set to the marked box mounted on the wall of the fourth floor hallway between buildings 8 and 16. Codes when relevant should be submitted on Course website.
4. The problem sets are due at noon on Friday the week after they were issued. There will be no extensions of deadlines for any problem sets in 20.320. **Late submissions will not be accepted.**

Please review the information about acceptable forms of collaboration, which is available on the course website and follow the guidelines carefully. Especially review the guidelines for collaboration on code. NO sharing of code is permitted.

1. TGFβ Receptor Signaling and Modeling Dynamic Systems (50 points)

In class we investigated the EGF Receptor and looked at its catalytic ability and role in cellular signaling. The cell has many receptors that share a similar relationship to signaling but ultimately lead to different downstream cellular responses. This problem asks you to look at a system conceptually similar to the EGF Receptor, derive qualitative and quantitative relationships among its relevant components and think about how you would explore this system experimentally to develop a computational model.

Part I – Modeling System Dynamics

A. There are two types of TGFβ receptors (TβR1 and TβR2) that come together to form a hetero-tetrameric complex before initiating downstream signaling. The stages in receptor activation are as follows: two TβR2 come together to form a homo-dimer. The homodimer binds TGFβ ligand. This complex recruits two TβR1 receptors to form the final hetero-tetrameric complex. Draw a cartoon to show the sequence of events. Label where the system would require reaction constants. You should start with "free" TβR1, TβR2, and TGFβ and end with the final TβR1- TβR2- TGFβ complex.

B. Now qualitatively sketch the amount of bound, signaling competent complex, the amount of receptor 1 and the amount of receptor 2 over time in the following scenarios (assume TGFβ ligand is in excess):
   a. You start with roughly equal concentrations of TβR1 and TβR2 ($[TβR1] = [TβR2]$)
   b. You have a 2-fold excess of TβR2 as compared to TβR1. ($2 \times [TβR1] = [TβR2]$)

C. Now write differential equations to describe your system. You should have a series of rate equations that include dynamics for free components in the system, the final tetrameric signaling complex and relevant intermediates (Six equations should be sufficient for this system).

Part II – Experimental Assays

D. What parameters would you need to fully characterize this system? What experimental assays might you use to determine these constants? Pick two and explain what information they give you, and what limitations come with using each approach.

E. As you may have noticed, in ODE systems, time-dependent parameters are important for characterizing dynamic systems. Yet, sometimes it's important to also perform equilibrium experiments to check your data. Being a good scientist, you perform an ITC experiment to extract K_D values for certain binding events in your system. You've done the experiment correctly, but the heat released data in your lab notebook is difficult to read.

Since you're running out of time to get your data published and presented, you wonder, if I try to make the experiment shorter by only doing one injection (i.e. adding all of the ligand into the system at once) – *is the heat released at the end of the experiment equivalent to the heat measured if you were to inject ligand through a series of injections? Defend your answer.*

F. What if you now wanted to take your understanding of the receptor system further to understand if the TGFβ receptor was involved in any pathway cross-talk – or look to see if it bound ligands other than TGFβ. Name two "discovery-driven" assays that would allow you to identify additional ligands for the receptor. Explain qualitatively what type of data these experiments might yield.

2. Michaelis-Menten Kinetics, Inhibitors and Application to Dynamic Systems (25 points)

A. State the two mathematical assumptions that go into deriving the Michaelis-Menten formalism and explain what these assumptions mean.

B. Using these two assumptions and the definition of K_M derive the Michaelis-Menten equation for rate of product formation.

C. What happens if we have a competitive inhibitor present? Write down the reactions for the system, and label all reactions with kinetic constants.

D. Assume the inhibitor, enzyme and inhibitor-enzyme complex are at equilibrium. The inhibitor equilibrium constant is K_I. Solve for $[dP]/dt$ change as a function of $[E_0]$, K_I, $[I]$, $[S]$ and K_m.

E. What chemical species is represented by $[I]$? Be precise. Several words is sufficient. Under what conditions can we reasonably estimate $[I]$?

F. The results suggest that adding an inhibitor is similar to modifying K_m. If we have 1 nM protein and 1uM inhibitor with K_I = 1uM, how does this change our apparent K_m?

G. Now that you've seen how Michaelis-Menten can be used to describe product formation in the presence of inhibitors, let's extend our previously-used TGFβ model to analyze the receptor's ability to function as an enzyme. As we saw with the EGF receptor, after complete complex formation, the TGFβ receptor system also acts as a kinase. However, in this case, the cytoplasmic segments of the TβR 2 receptor phosphorylate serine and threonine residues on the TβR1 receptor. What additional equation would you use to describe the rate of phosphorylated TβR1 receptor? (For this example, you can ignore the fact that multiple sites on TβR1 will be phosphorylated).

**3. Protein Binding Arrays and Extraction of Equilibrium Parameters
(25 points)**

In class we discussed protein binding arrays and how they might be used to test for the binding interactions of many proteins in a high-throughput fashion. In both of our EGF and TGFβ systems, once these receptors become phosphorylated, there is a series of interacting proteins that can bind phosphorylated peptides. In this case you are interested in looking at the binding of SARA (Smad anchor for receptor activation) to phosphorylated sites on the TβR1 receptor before downstream signaling is initiated. The SARA proteins function to recruit Smad proteins to the receptor and allow for subsequent phosphorylation events.

You have obtained ligand binding data of fluorescently-labeled SARA protein to various phosphosites on TβR1. The data is in the file SARAbinding.xls. Each row corresponds to a different phosphosite (pS10, pT27, pT30, pS56, pT60, pT65, pS73) and each column corresponds to a different starting ligand concentration in nM (1   5   10   20   40   80   100   150   200).

   A. Load the data into MATLAB, plot the raw data. Then, using nlinfit, determine the K_D values for each binding interaction. Assume that you are operating under PFOA conditions. Note: you will have to convert the data into fractional saturation and each curve will have its own equilibrium constant.

   B. If we assume affinity is a good metric for a more active adapter protein, which site is the most important for eliciting a SARA-induced signaling response? What is wrong with assuming that K_d is a good proxy for the likelihood of a particular signaling event occurring?

   C. After calculating the K_D replot "theoretical" fractional saturation curves on the same plots with your experimental data. For which phosphosites were you able to fit better? Which phosphosites is your fitted curve less correct?

   D. To be more quantitative in your analysis of fit, now plot the residual values between the model and the data. What can you conclude about the appropriateness of your model? (Consider normalizing the residuals so you can compare across sites!) Using the residuals plot and the plot from part C, can you hypothesize any biological behavior that our model equation ignores?

20.320 Analysis of Biomolecular and Cellular Systems
Fall 2012