# 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 Course website. 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.

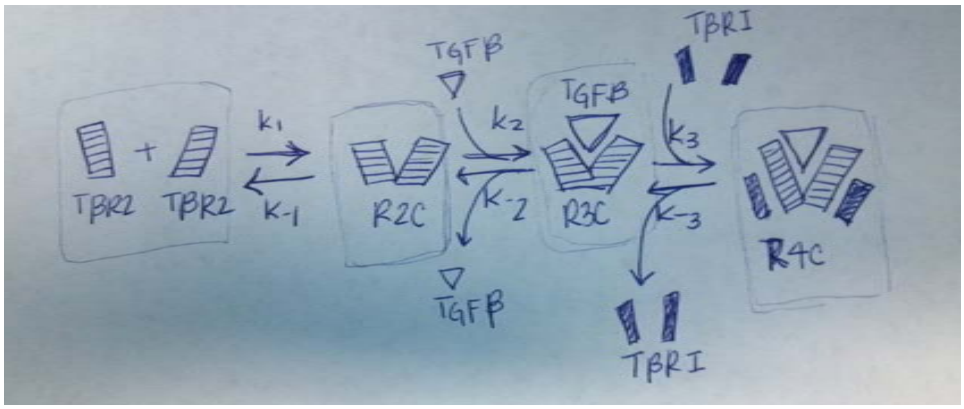
## 1. TGFβ Receptor Signaling and Modeling Dynamic Systems (50 points)

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. You should start with "free" TβR1, TβR2, and TGFβ and end with the final TβR1-TβR2- TGFβ

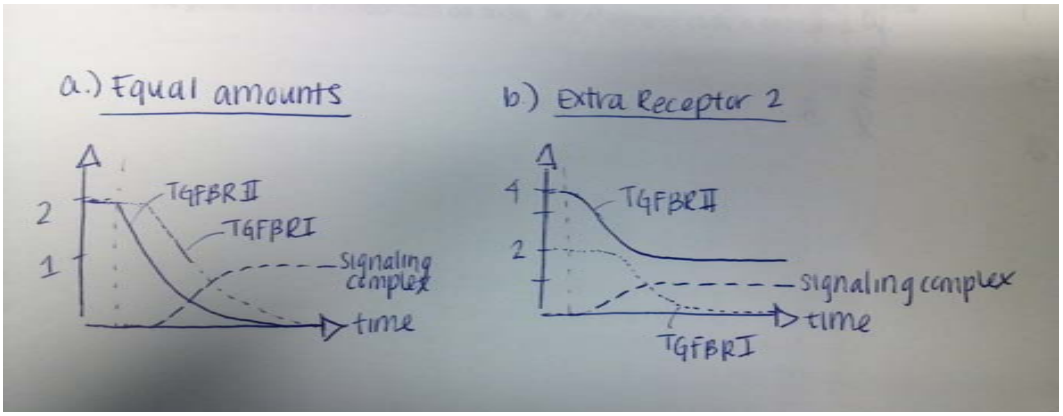
5




points for including all relevant species
2 points for including rate constants on reactions
1point for correct stoichiometry

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 x [TβR1] = [TβR2]*)

10 points (5 points per graph)




**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.

12 points – 2 for each correct differential equation with relevant rate constants and species

$$\frac{d[TR2]}{dt} = -k_1[TR2]^2 + k_{-1}2[R2C]$$

$$\frac{d[R2C]}{dt} = k_1[TR2]^2 - k_{-1}[R2C] - k_2[R2C][TGFb] + k_{-2}[R3C]$$

$$\frac{d[TGFb]}{dt} = -k_2[R2C][TGFb]^2 + k_{-2}2[R3C]$$

$$\frac{d[R3C]}{dt} = k_2[R2C][TGFb]^2 - k_{-2}[R3C] - k_3[R3C][R1]^2 + k_{-3}[R4C]$$

$$\frac{d[R1]}{dt} = -k_3[R3C][R1]^2 + k_{-3}2[R4C]$$

$$\frac{d[R4C]}{dt} = k_3[R3C][R1]^2 - k_{-3}[R4C]$$

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 experiments and explain what information they give you, and what limitations come with using each approach.

9 points (3 points for identifying all of the constants, and 3 for each appropriately selected and explained experimental assay)

You would need 6 rate constants – the on and off rates for the formation of the receptor 2 complex, the TGFβ-bound complex and the final tetrameric complex.

The experiments selected need to be able to determine rate constants (equilibrium experiments are not encouraged, but could be used if they extract either an on or off rate from a dynamic experiment and then use an equilibrium experiment to extra the other on or off rate from the Kd).

Some examples:
SPR – gives you on and off rates, limited by the ability to attach a binding partner to the surface, the reactions are slow. This might be difficult to measure binding to large complexes (since you have many components coming together on the surface).
FRET – could be adapted to give parameters, limited by the fact that it's very low throughput, each experiment has a low yield, only measures proximity, not actual interaction

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.*

5 points
The heat evolved at the end of the experiment is the same as the total heat evolved over all of the injections because ITC is a state experiment.

Only the final endpoints matter, not how you get there.

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.

6 points (3 points for each assay selected and an appropriate explanation of the type of data and the limitations)

Co-IP – will identify things that are in complex with your protein of interest when the sample has been cross-linked. Usually coupled with mass spec, so this would yielt sequenced peptide data.

Yeast-2-Hybrid – Again will identify anything capable of binding your protein of interest and capable of being expressed in the nucleus. The data is expressed as a fluorescent read out per combination of binding partners.

Protein Array – A fluorescent-based experiment that gives you a read out of which peptides bind your protein of interest.

FRET and PLA are not appropriate here because you want to discover novel binders and don't want to look at binding dynamics for things that you already know interact.

**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. (1 point)

d[ES]/dt = 0

$k_{-2} = 0$

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

(7 points)

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] - k_2[ES] = 0$$

$$\frac{[E][S]}{[ES]} = \frac{k_{-1} + k_2}{k_1} = K_m$$

By mass conservation, $[E] = [E_0] - [ES]$, so

$$\frac{([E_0] - [ES])[S]}{[ES]} = K_m$$

$$[ES] = \frac{[E_0][S]}{K_m + [S]}$$

$$\frac{d[P]}{dt} = k_2[ES] = \frac{k_2[E_0][S]}{K_m + [S]} = \frac{V_{max}[S]}{K_m + [S]}$$

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

(3 points)

$$\mathrm{E{\cdot}I} + \mathrm{S} \underset{k_{on}}{\overset{k_{off}}{\rightleftharpoons}} \mathrm{E} + \mathrm{I} + \mathrm{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{E{\cdot}S} + \mathrm{I} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{E} + \mathrm{P} + \mathrm{I}$$

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 $K_m$.

(7 points)

$$\begin{aligned}
\frac{d[P]}{dt} &= k_2[E \cdot S] \\
&= \frac{k_2[E \cdot S][E_{total}]}{[E] + [E \cdot S] + [E \cdot I]} \\
&= \frac{k_2 \frac{[E][S]}{K_m}[E_{total}]}{[E] + \frac{[E][S]}{K_m} + \frac{[E][I]}{K_i}} \\
&= \frac{k_2[S][E_{total}]}{K_m + [S] + \frac{[I]K_m}{K_i}} \\
&= \frac{k_2[S][E_{total}]}{K_m\left(1 + \frac{[I]}{K_i}\right) + [S]}
\end{aligned}$$

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

(2 points)
[I] is the concentration of the UNBOUND inhibitor. [I] ~ $[I_0]$ when the concentration of inhibitor is far greater than the concentration of enzyme.

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 $K_m$?

(2 points)
It doubles it!

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 points)

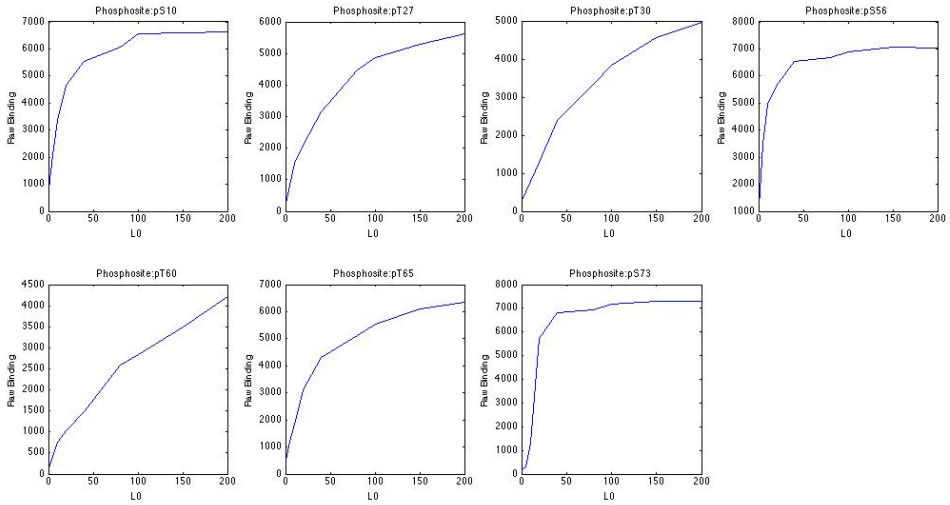
$$\frac{d[R4C^*]}{dt} = \frac{V_{\max}[R4C]}{K_M + [R4C]}$$ where R4C* is the activated complex

## 3. Protein Binding Arrays and Extraction of Equilibrium Parameters

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.

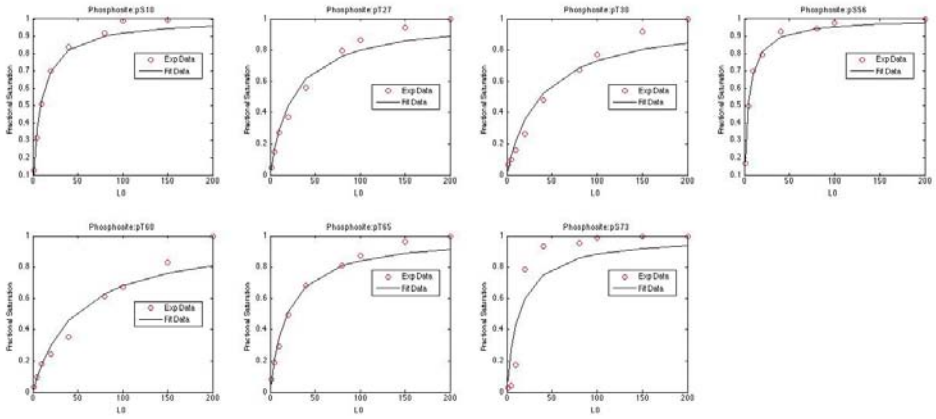



Kd values in nM
8.7519; 24.8579; 36.3450; 4.6511; 47.0911; 18.9401; 13.3013
7 points (4 for plotting the raw data and 3 for getting the right values)

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?

(4 points) It appears that the pS at position 56 has the best Kd value. Having a favorable Kd value doesn't take into account the cellular context or relative protein concentrations at the membrane and so it may not be a good estimate for the likelihood of a particular signaling response occurring.

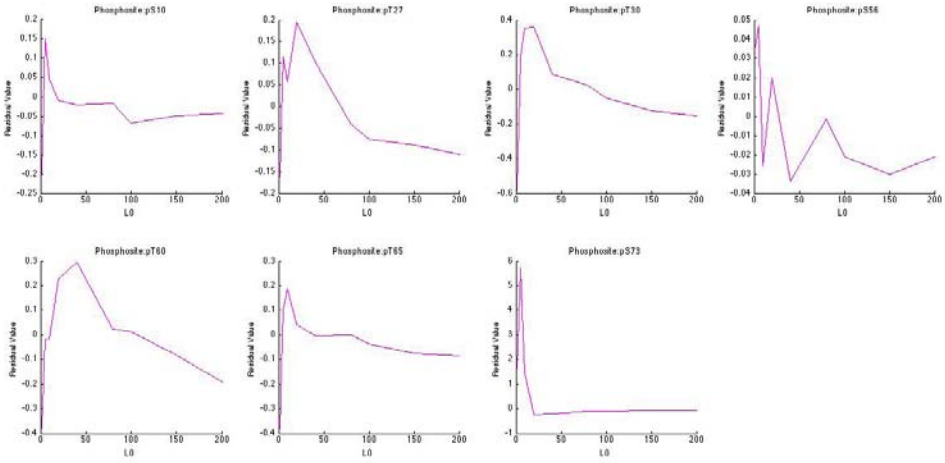
C. After calculating the $K_D$ replot "theoretical" fractional saturation curves on the same plots with your experimental data. For which phosphosites were you able to fit better? Which phosphosites is your fitted curve less correct?




7 points (4 for the plot and 3 for the explanation)
It appears that most sites are well fit with the exception of pT30 and pS73 aren't as well fit with this model.

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 using ignores?

7 points (4 for the plot and 3 for the explanation)
Overall the residual values are small so it seems that the model is overall well fit (with the exception of site 73). It appears that pS73 might be undergoing some sort of cooperative binding. You can see this a little in the steepness of the binding curve from part C.


Phosphosite:pS10
Phosphosite:pT27
Phosphosite:pT30
Phosphosite:pS56
Phosphosite:pT60
Phosphosite:pT65
Phosphosite:pS73
Residual Value
L0

20.320 Analysis of Biomolecular and Cellular Systems
Fall 2012