

# Cellular Bioengineering Project 1 Solutions

Alisha Sarang-Sieminski

February 18, 2014

## I. Part 1 (5 points)

As described in the problem statement, gp120 is present on the surface of HIV. It binds to CD4 on the surface of T-cells. After the initial binding event, CCR5 (also on the T-cell surface) also binds to gp120. We are interested in describing these binding events using first-order rate equations for receptor-ligand binding. A schematic of the binding is shown in Figure 1A. Forward and reverse rate constants for gp-120 ( $R$ ) binding to CD4 ( $L_1$ ) to form the first complex,  $C_1$ , are  $k_1$  and  $k_2$ , respectively. The forward and reverse rate constants for  $C_1$  binding to CCR5 ( $L_2$ ) to form the ternary complex,  $C_2$ , are  $k_3$  and  $k_4$ , respectively.

The rate equations describing these interactions are as follows:

$$\frac{dC_1}{dt} = k_1RL_1 - k_2C_1 - k_3C_1L_2 + k_4C_2$$

$$\frac{dC_2}{dt} = k_3C_1L_2 - k_4C_2$$

$$\frac{dL_1}{dt} = -k_1RL_1 + k_2C_1$$

$$\frac{dL_2}{dt} = -k_3C_1L_2 + k_4C_2$$

$$\frac{dR}{dt} = -k_1RL_1 + k_2C_1$$

For this problem, since gp120, CD4, and CCR5 are on a cell surface, I will explicitly choose units of numbers per cell as opposed to the units of [M] we have been using for ligands. Thus, the mass balance equations are:

$$L_{1,0} = L_1 + C_1 + C_2$$

$$L_{2,0} = L_2 + C_2$$

$$R_T = R + C_1 + C_2$$

where  $L_{1,0}$ ,  $L_{2,0}$ , and  $R_T$  are the total number of CD4, CCR5, and gp120 molecules per cell, respectively.

Note that you could also chose to express CD4 and CCR5 in [M], in which case  $L_{1,0} = L_1 + \frac{n}{N_A}C_1 + \frac{n}{N_A}C_2$  and  $L_{2,0} = L_2 + \frac{n}{N_A}C_2$ , where  $L_{1,0}$  and  $L_{2,0}$  are the initial concentrations of CD4 and CCR5, respectively, in [M].  $n$  is the number of cells per volume, and  $N_A$  is Avogadro's number (number of molecules per mole).

Assuming excess ligand gives the approximations,  $L_1 \sim L_{1,0}$  and  $L_2 \sim L_{2,0}$ . Note that, unlike the examples involving soluble ligands we've discussed in class, this implies that there is an excess of cell-surface proteins available for binding relative to the number of complexes. This is probably still a fine assumption given that the number of HIV particles binding to the cell at a given time is probably small.

At steady state, all of the time differentials are zero.

Rearranging to eliminate  $R$  and  $C_1$ , we can obtain an expression for full complexes,  $C_2$ , as a function of system inputs and parameters:

$$C_2 = \frac{R_T L_{1,0} L_{2,0}}{K_{D2}(L_{1,0} + K_{D1}) + L_{1,0} L_{2,0}}$$

where  $K_{D4} = \frac{k_2}{k_1}$  and  $K_{D5} = \frac{k_4}{k_3}$ . It is the custom that the rate constants are denoted with lower case 'k' and the equilibrium constants are denoted with upper case 'K'. Also, note that there are many other algebraic rearrangements of this that are also correct.

It's a little hard to intuit the behavior of this equation. Thus, we could also rearrange to express full complexes as:

$$C_2 = \frac{R_T L_{2,0}}{K_{D2}(1 + \frac{K_{D1}}{L_{1,0}}) + L_{2,0}}$$

Getting the expression for  $C_2$  in this particular (rather familiar) form shows us that a plot of full complexes as a function of CCR5 concentration, or  $L_{2,0}$ , will have the same  $R_T$ , regardless of CD4 concentration, or  $L_{1,0}$ . This highlights the importance of really looking at your equations and understanding them because, for example, just looking at the graph at low concentrations of CD4 might lead you to believe that  $R_T$  appears to vary with CCR5. Further, this form highlights an apparent  $K_D$ ,  $K_{D,app} = K_{D2}(1 + \frac{K_{D1}}{L_{1,0}})$ . Graphs showing the effect of varying  $K_{D2}$  (Fig. 1B) or  $L_{2,0}$  (Fig. 1C) are shown in the figures at the end of this document.

In grading this section I looked at your well-labeled schematics (1 point), your equations (1 point), a clear statement of assumptions and their meanings and indication of the units of the different parameters (1 point), and getting the right solution (1 point). For the graphs (1 point), I was looking at whether you varied some parameters and drew some conclusions from the graphs.

## II. Part 2 (7 points)

For simplicity, we only consider the binding of gp120 ( $R$ ) to CD4 ( $L_1$ ) in part 2. We also introduce the inhibitor BMS-378806 ( $I$ ). We first consider competitive binding where CD4 and BMS-378806 can bind to gp120 in the same site. The forward and reverse reactions between gp120 and CD4 are again given rate constants  $k_1$  and  $k_2$ , respectively. Let the forward and reverse rate constants for gp120 and BMS-378806

be  $k_5$  and  $k_6$ , respectively. The equilibrium constants are  $K_{D1} = \frac{k_2}{k_1}$  and  $K_I = \frac{k_6}{k_5}$ . The gp120-CD4 complex is still  $C_1$  and the gp-120-BMS complex is  $C_3$ . Schematics are shown in Figure 2A.

The rate equations defining this system are:

$$\frac{dC_1}{dt} = k_1RL_1 - k_2C_1$$

$$\frac{dC_3}{dt} = k_5RI - k_6C_3$$

$$\frac{dL_1}{dt} = -k_1RL_1 + k_2C_1$$

$$\frac{dI}{dt} = -k_5RI + k_6C_3$$

$$\frac{dR}{dt} = -k_1RL_1 + k_2C_1 - k_5RI + k_6C_3$$

Again, the units of our various components could be anything, so long as we are explicit about them. Given that the experiment done in this case involves soluble CD4 and BMS binding to surface bound gp120 in a well, I have chosen  $R$ ,  $C_1$ , and  $C_3$  to be expressed as molecules per well and  $L_1$  and  $I$  to be expressed in [M]. Thus our mass balances are:

$$L_{1,0} = L_1 + \frac{C_1}{N_A V_w}$$

$$R_T = R + C_1 + C_2$$

$$I_0 = I + \frac{C_3}{N_A V_w}$$

where  $R_T$  is the total number of gp120 molecules per well,  $L_{1,0}$  is the initial concentration [M] of CD4,  $I_0$  is the initial concentration [M] of BMS,  $V_w$  is volume of liquid per well, and  $N_A$  is Avogadro's number.

Once again, we will assume excess ligand ( $L_1 \sim L_{1,0}$  and  $I \sim I_0$ ) and steady state (time differentials are zero). Solving for  $C_1$  in terms of things we know or can measure gives:

$$C_1 = \frac{R_T L_{1,0}}{K_{D1} \left(1 + \frac{I_0}{K_I}\right) + L_{1,0}}$$

Taking this a step further, we can say that:

$$C_1 = \frac{R_T L_{1,0}}{K_{D1,app} + L_{1,0}}$$

where  $K_{D1,app} = K_{D1} \left(1 + \frac{I_0}{K_I}\right)$ .

The point is that you know the general form that you are looking for ( $C = \frac{R_T L_0}{K_D + L_0}$ ), so you should try to rearrange to get something that looks like that to make meaning out of the equations and the data. Many

people got to  $C_1 = \frac{R_T L_{1,0} K_{D2}}{L_{1,0} K_{D2} + I K_{D1} + K_{D1} K_{D2}}$ , and stopped. Dividing by  $K_{D2}$  and rearranging gives the more intuitive form to work with.

Expressing the equation this way lets us see that in the presence of inhibitor, the number of CD4-gp120 complexes will appear to have a different  $K_{D1}$ , but an unchanged  $R_T$ . You should have a check in that for  $I_0 = 0$  we get back the expression for single R-L equilibrium binding,  $C_1 = \frac{R_T L_{1,0}}{K_{D1} + L_{1,0}}$ . It also shows us that  $K_{D1,app}$  should linearly increase with  $I$ !

Now we can get to business plotting the given data on a Scatchard plot. Rearranging, we get:

$$\frac{C_1}{L_{1,0}} = \frac{R_T}{K_{D1,app}} - \frac{L_{1,0}}{K_{D1,app}}$$

Plotting the data on a Scatchard plot (Fig. 2B) and fitting for the slope and intercept, we obtain:

[BMS] (nM)	0	0.8	1.6	3.2
slope	-0.97	-0.27	-0.16	-0.08
intercept	1.32	0.37	0.21	0.12
$K_{D,app}$ (nM)	1.03	3.75	6.30	11.82
$R_T$ (abs)	1.35	1.37	1.35	1.37

We can get  $K_{D1} = 1.03$  nM from the BMS = 0 set and, averaging the  $R_T$  values for all 4 lines, gives us that  $R_T = 1.36$  in absorbance units! Further, if we plot  $K_{D,app}$  vs.  $I$  (Fig. 2C), we see that it is a line with a slope of 3.37. Thus, our data is consistent with a competitive binding mechanism and we can even solve for  $K_I = 0.31$  nM! Amazing, no?

For this section, I was really looking for you to really use your model along with the data to make this type of conclusion and really solve for the parameters because you can. The schematics (0.5 points), equations (1 point), and assumptions (0.5 points) were rather simple. A correct solution was good (1 point) but inspecting to see that you could turn it into a  $K_{D,app}$  (in some way) was even better (1 point). The focus was on using the data to make a Scatchard plot (1 point) and solving for  $R_T$ ,  $K_{D1}$ , and  $K_I$  (1.5 points) and giving them the correct units. Finally, correctly explaining why the data was consistent with competitive binding was looked at (0.5 points).

### III. Part 3 (3 points)

For part 3, the question was rather ambiguous as to whether the new inhibitor bound to gp120 or CCR5. The equations for inhibitor ( $I$ ) binding to gp120( $R$ ) and changing the affinity for CCR5 ( $L_2$ ) binding to the gp120-CD4 complex ( $C_1$ ) are given. Versions that explicitly pointed out that binding the inhibitor after the gp120-CD4-CCR5 complex formed would not help, were also considered correct. Naming conventions the same as in Part 1. Schematics are shown in Figure 3.

Let  $C_3$  be inhibitor bound to gp120-CD4 ( $C_1$ ) (with rate constants for this interaction,  $k_5$  and  $k_6$ ) and  $C_4$  be the quaternary complex of gp120-CD4-CCR5-inhibitor. The rate constants for binding of gp120-CD4-inhibitor ( $C_3$ ) to CCR5 ( $L_2$ ) are  $k_7$  and  $k_8$  and those for binding of inhibitor to gp120-CD4-CCR5 ( $C_2$ ) are  $k_9$  and  $k_{10}$  (pew!). All of this mess gives us the equations:

$$\frac{dC_1}{dt} = k_1 R L_1 - k_2 C_1 - k_3 C_1 L_2 + k_4 C_2 - k_5 C_1 I + k_6 C_3$$

$$\frac{dC_2}{dt} = k_3C_1L_2 - k_4C_2 - k_9C_2I + k_{10}C_4$$

$$\frac{dC_3}{dt} = k_5C_1I - k_6C_3 - k_7C_3L_2 + k_8C_4$$

$$\frac{dC_4}{dt} = k_7C_3L_2 - k_8C_4 + k_9C_2I - k_{10}C_4$$

$$\frac{dL_1}{dt} = -k_1RL_1 + k_2C_1$$

$$\frac{dL_2}{dt} = -k_3C_1L_2 + k_4C_2 - k_7C_3L_2 + k_8C_4$$

$$\frac{dI}{dt} = -k_5C_1I + k_6C_3 - k_9C_2I + k_{10}C_4$$

$$\frac{dR}{dt} = -k_1RL_1 + k_2C_1$$

And we can write the epic mass balance equations for all of our many friends here:

$$L_{1,0} = L_1 + C_1 + C_2 + C_3 + C_4$$

$$L_{2,0} = L_2 + C_2 + C_4$$

$$R_T = R + C_1 + C_2 + C_3 + C_4$$

$$I_0 = I + \frac{n}{N_A}(C_3 + C_4)$$

where  $L_{1,0}$ ,  $L_{2,0}$ , and  $R_T$  are the total number of CD4, CCR5, and gp120 molecules per cell, respectively, and  $I_0$  is the molar concentration of inhibitor.

Of course, we don't want to solve this, but the general idea is that we want an expression for  $(C_2 + C_4)$  since this is ultimately the functional unit for HIV-T-cell binding. This expression should eliminate  $R$ ,  $C_1$ , and  $C_3$  and will use the excess ligand assumption.

For this final bit, I was looking to make sure you could do the book keeping in this type of model when things get hairy – basically trying to force you to use the pictures because you can't just do this in your head. I looked at schematics (1 point), equations (1 point), and some description of how you would solve (1 point).

Figure 1

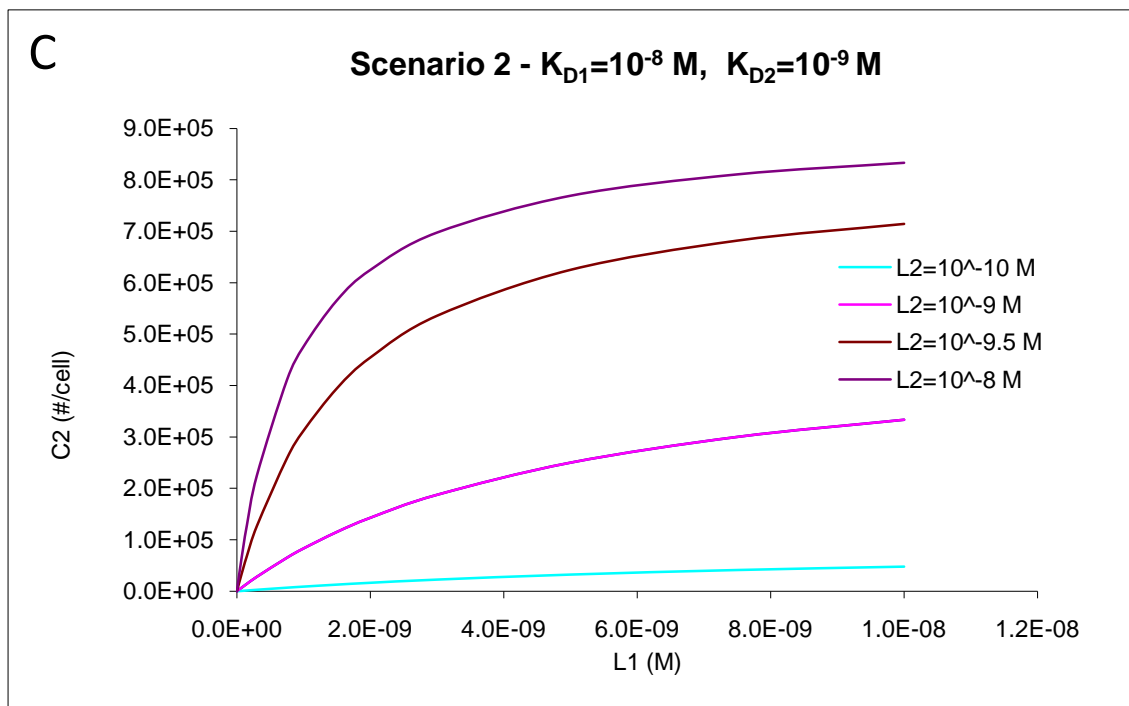
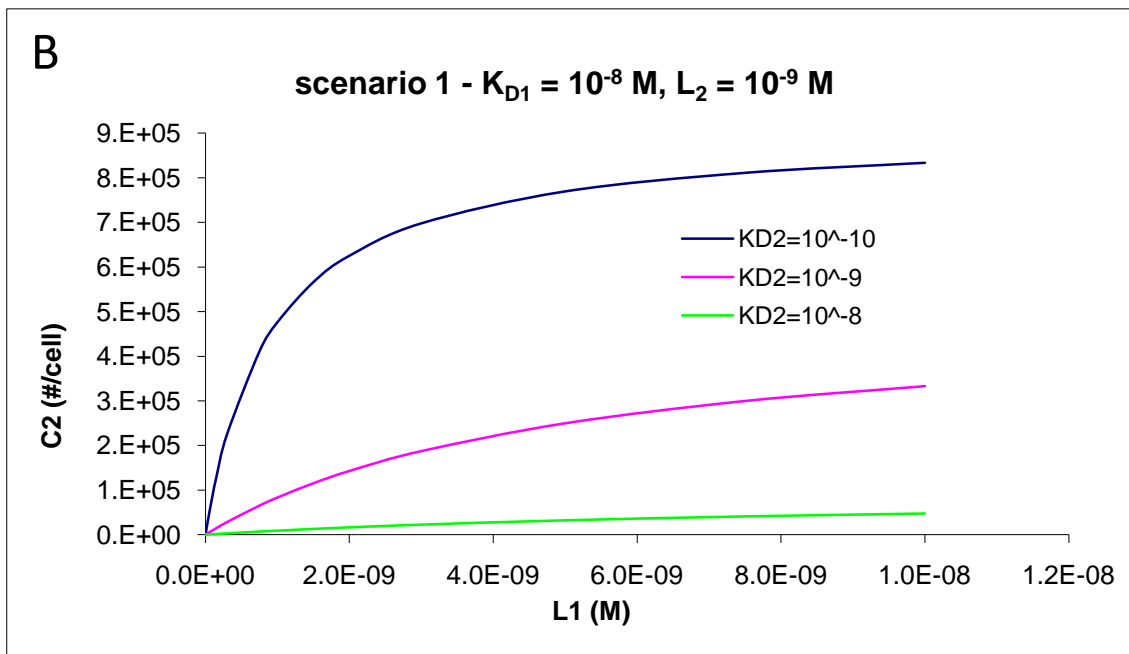
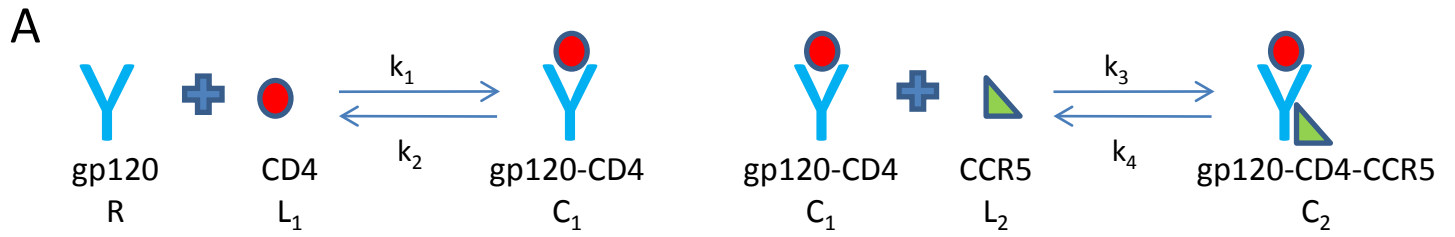
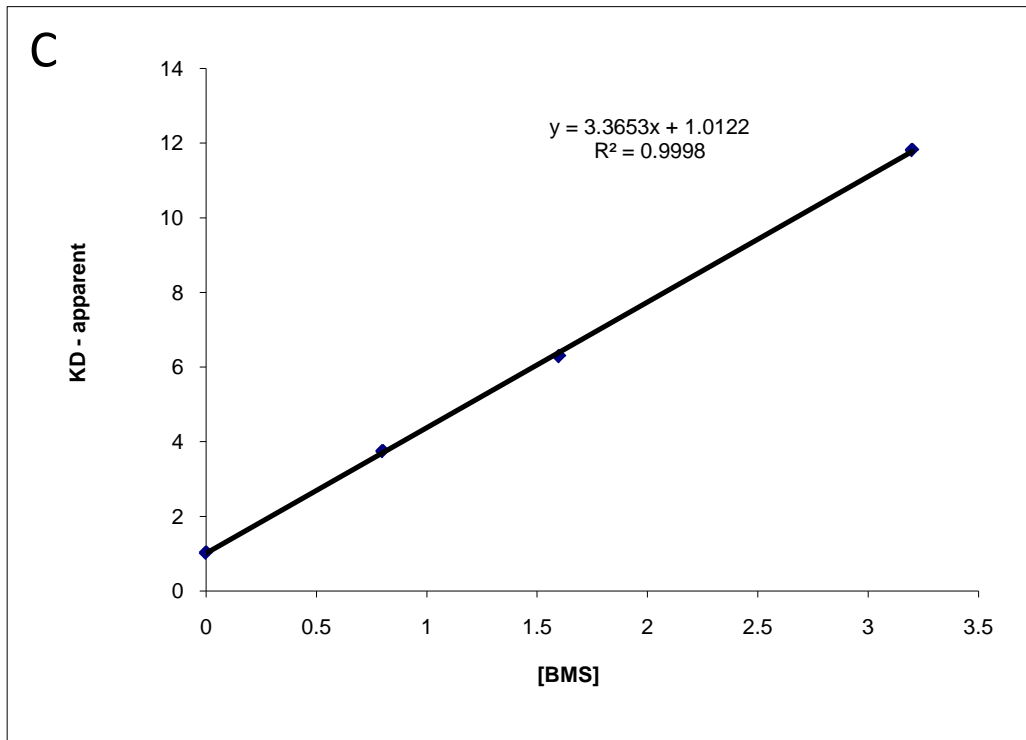
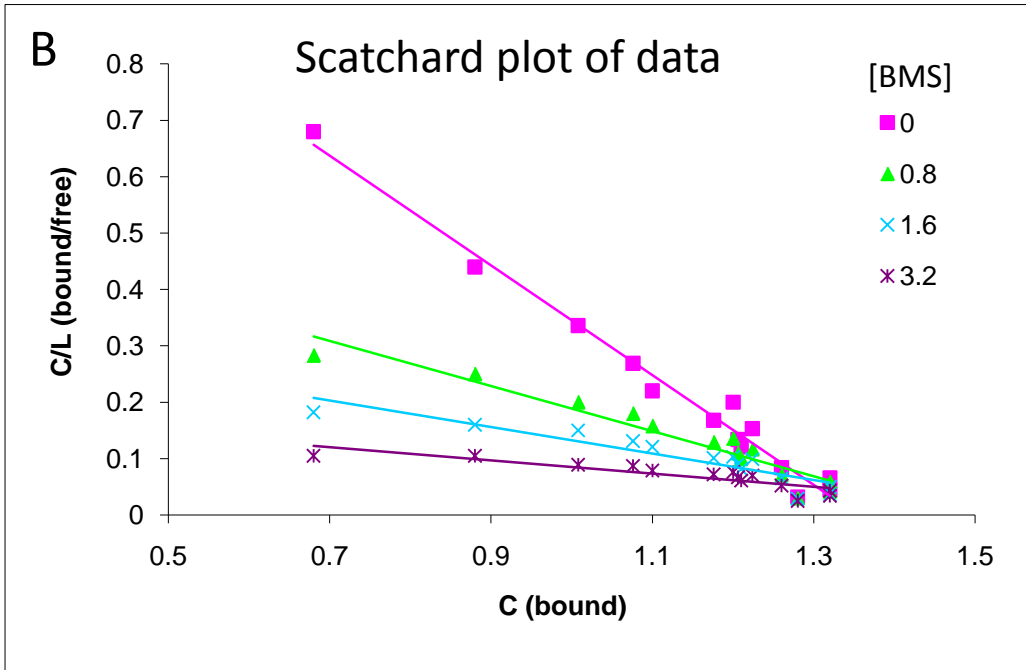
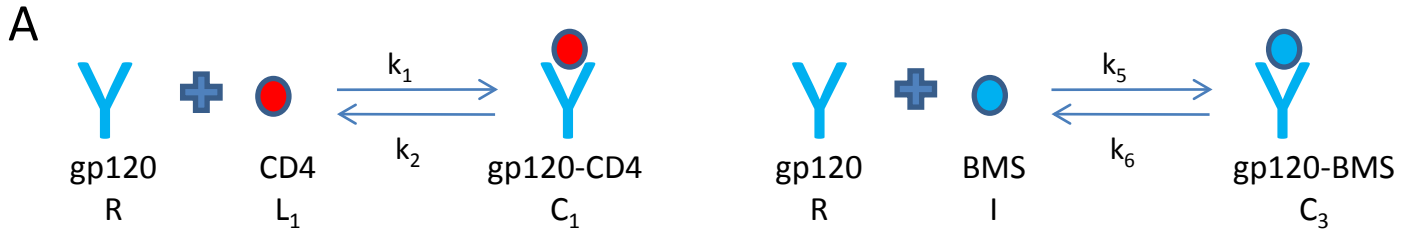


Figure 2



**Figure 3**

