

Due: 3/13/06

1. 20mM NaCl:  $k_{on}^{I_1} = 2 \times 10^6 \frac{L}{mol \cdot s}$   
 50mM NaCl:  $k_{on}^{I_2} = 1.12 \times 10^6 \frac{L}{mol \cdot s}$

(a)  $\log k_{on}^{I_1} = \log k_{on}^{I=0} + 1.02 z_p z_L \sqrt{I_1}$   
 $I_1 = \frac{1}{2} \sum_i m_i z_i^2 = \frac{1}{2} \left( \frac{20 \text{ mmol}}{\text{kg}} (1)^2 + \frac{20 \text{ mmol}}{\text{kg}} (-1)^2 \right) \Rightarrow \sqrt{I_1} = 0.141$   
 $\log k_{on}^{I_2} = \log k_{on}^{I=0} + 1.02 z_p z_L \sqrt{I_2} \Rightarrow I_2 = \frac{1}{2} \sum_i m_i z_i^2, m_i = \frac{50 \text{ mmol}}{\text{kg}}$   
 $\sqrt{I_2} = 0.224$

$\log k_{on}^{I_1} = \log k_{on}^{I=0} + 0.144 z_p z_L$  Eqn ①

$\log k_{on}^{I_2} = \log k_{on}^{I=0} + 0.228 z_p z_L$  Eqn ②

Multiply ① by 1.58, subtract Eqn ②:

$1.58 (\log k_{on}^{I_1} - \log k_{on}^{I_2}) = 0.58 \log k_{on}^{I=0} \Rightarrow k_{on}^{I=0} = 5.42 \times 10^6 \frac{L}{mol \cdot s}$

(b)  $\log k_{on}^{I_2} = \log k_{on}^{I=0} + 1.02 z_p z_L \sqrt{I_2} \Rightarrow z_p z_L = -3$

It is not possible to determine individual charge on the proteins. We know only that  $z_p z_L = -3$  (where  $z_p, z_L$  are the valence charges on the 2 interacting proteins). Additionally, we know that the 2 proteins are oppositely charged. ( $z_p z_L$  is negative,  $k_{on}$  decreases with increasing NaCl, and NaCl shields interactions b/t oppositely charged proteins).

(c)  $\log k_{on}^{I=150} = \log k_{on}^{I=0} + 1.02 z_p z_L \sqrt{I_{CaCl_2}}$  \*Note: moles of Cl = 2 x moles of CaCl<sub>2</sub>  
 $I_{CaCl_2} = \frac{1}{2} \left( \frac{150 \text{ mmol}}{\text{kg}} (2)^2 + \frac{300 \text{ mmol}}{\text{kg}} (-1)^2 \right) \Rightarrow \sqrt{I} = 0.67$

$\log k_{on}^{I=150} = 4.681 \Rightarrow k_{on} = 4.8 \times 10^4 \frac{L}{mol \cdot s}$

protein-protein interactions VERY heavily shielded by ions in solution.

\* NOTE a/b molality:  $molality = \frac{\text{moles of solute}}{\text{kg of solvent}}$

ex. 50 mM CaCl<sub>2</sub> =  $\frac{50 \text{ mmol solute}}{\text{L of solvent}}$ , assume solvent = H<sub>2</sub>O w/ density 1 kg/L  
 molality of CaCl<sub>2</sub> =  $\frac{50 \text{ mmol}}{\text{kg}}$ , BUT molality of Cl<sup>-</sup> is  $\frac{2(50) \text{ mmol}}{\text{kg}}$

2.  $K_i = \frac{\binom{n}{i-1}}{\binom{n}{i}} K^M$       4th ligand:  $K_4 = \frac{\binom{6}{3}}{\binom{6}{4}} K^M = \frac{3! 3!}{4! 2!} K^M = \frac{4}{3} K^M$

Assume all sites independent & equal

3rd ligand:  $K_3 = \frac{\binom{6}{2}}{\binom{6}{3}} K^M = \frac{3}{4} K^M$

$\frac{K_4}{K_3} = \frac{\frac{4}{3}}{\frac{3}{4}} = \frac{16}{9} \Rightarrow$  binding 4th ligand has 1.78 times lower affinity than binding 3rd ligand

3.  $\Delta G_{1,2} = -2.1 \frac{\text{kcal}}{\text{mol}} = \Delta G_2 - \Delta G_1$

(a)  $\Delta G_i = RT \ln K_i + RT \ln \left( \frac{\binom{n}{i}}{\binom{n}{i-1}} \right)$

$\Delta G_{i,j} = RT \ln \left( \frac{K_j}{K_i} \right) + RT \ln \left( \frac{\binom{n}{j} \binom{n}{i-1}}{\binom{n}{j-1} \binom{n}{i}} \right) \Rightarrow \frac{K_j}{K_i} = \exp \left( \frac{\Delta G_{i,j}}{RT} - \ln \left( \frac{\binom{n}{j} \binom{n}{i-1}}{\binom{n}{j-1} \binom{n}{i}} \right) \right)$

$\Rightarrow = \exp \left[ \frac{-2.1 \text{ kcal/mol}}{RT} - \ln \left( \frac{\binom{4}{2} \binom{4}{0}}{\binom{4}{1} \binom{4}{1}} \right) \right]$        $R = 1.9872 \text{ cal}$   
 Assume  $T = 300\text{K} \Rightarrow \frac{K_2}{K_1} = 0.078$

(b) independent sites:  $K_2 = \frac{\binom{4}{1}}{\binom{4}{2}} K^M$        $K_1 = \frac{\binom{4}{0}}{\binom{4}{1}} K^M$

$\frac{K_2}{K_1} = \frac{\frac{4}{6}}{\frac{1}{4}} = \frac{16}{6} = \frac{8}{3}$

You can also solve by setting  $\Delta G_{1,2} = 0$ , and you'll get this same ratio

4. Let's do order of magnitude estimates, assume  $T = 300\text{K}$

Eyring:  $k_{\text{on}}^{\text{max}}$  given by frequency of vibrations  $\Rightarrow k_{\text{on,Ey}}^{\text{max}} = \frac{k_B T}{h} e^{-\frac{\Delta S^\ddagger}{R}} \approx \left[ \frac{k_B T}{h} \approx 10^{13} \frac{\text{L}}{\text{mol}\cdot\text{s}} \right]$

collision theory:  $k_{\text{on}}^{\text{max}}$  given by when  $E_a \rightarrow 0$ , thus  $k_{\text{on}}^{\text{max}}$  also relates to the freq. factor

$k_{\text{on}}^{\text{max}} = A = N_{\text{avg}} \pi \sigma_{\text{AB}}^2 \sqrt{\frac{8k_B T}{\pi \mu}} \Rightarrow \sigma_{\text{AB}} \approx 10^{-9} \text{ m}$   
 $\mu \approx 10^2 \text{ kDa} \approx 10^{-22} \text{ kg} \Rightarrow \sqrt{\frac{8k_B T}{\pi \mu}} \approx 10^7 \frac{\text{m}}{\text{s}}$

order of magnitude approx:  $k_{\text{on,collision}}^{\text{max}} \approx 10^7 \frac{\text{L}}{\text{mol}\cdot\text{s}}$

Eyring theory has much higher max rate than collision theory given this estimate. One explanation is that collision theory does not take into account the effects of orientation, which could significantly help or hinder rate of binding.

Solution to problem 5.

Fractional saturation is most frequently defined as  $y = \text{fraction of binding sites that are occupied (number of sites bound/total number of sites)}$ . In this problem, the data was created using this definition. In class we called  $v = \text{“mol ligand bound”/“mol protein bound”}$  the fractional saturation as well. You have also seen it defined as the fraction of total protein that is in complex. In fact, we accidentally use both  $y$  and  $v$  in different parts of this problem, so we will be more careful in the future about defining this term each time we use it! And most answers will be accepted as solutions. Please go through the solution in detail though, and if you'd like you can try creating data using the definition for  $v$  and fitting data.

(a)  $v = \text{mol ligand bound / mol protein bound}$  in this case of independent binding sites is:  
$$n[L]/([L]+K_{\mu})$$

This was derived in class on 3/8/06 and is given as equation 3.30 in the W&T text.

MATLAB code is attached that fits the variable  $n$  to the given data. The fit value of  $n$  is about 1.17 and from your residual plot you can see that the fit is quite poor. Note that since we fit the equation  $v$  to data created using the equation  $y$ , this is not so surprising. But even if the data had been created using the cooperative equation  $v$  below, the fit would look similarly bad!

(b)  $v = \text{mol ligand bound / mol protein bound}$  in the fully cooperative case is:

$$n[L]^n/([L]^n + Kd)$$

And  $y = \text{fraction of occupied binding sites}$  is:

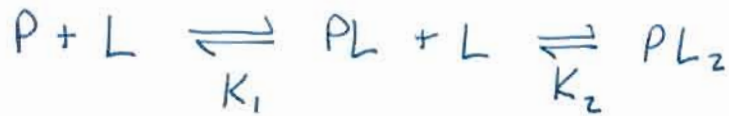
$[L]^n/([L]^n + Kd)$ , where  $Kd = K_{\mu}^n$  in both equations. Since the data was created using this second equation, we fit the second equation to the data in the code. However, you would just make a small change to `Y_cooperative.m` to fit the first equation.

$Kd$  is the macroscopic  $Kd$  of binding 4 ligands to the receptor at once, and  $K_{\mu}$  is the  $Kd$  of binding just one ligand to one binding site, which was the constant given in the problem. Here,  $Kd = K_{\mu}^n$  because the binding is fully cooperative. Please see the attached page of notes for an explanation of this equality. Here, the fit value of  $n$  is about 3.2 and the residuals from the fit look very random. Therefore, we assume that the binding sites are cooperative.

(c) We assumed in (b) that the binding sites are *fully* cooperative. This means that once one ligand binds, all the other binding sites are immediately filled as well. So for 4 binding sites, either 0 ligands bind or 4 ligands bind. This is not a very realistic scenario, since for a real protein, one ligand binding may change the protein structure so that the affinity is greater for the next ligand, but it will not likely force the next ligand to bind immediately. This unrealistic assumption is still useful for us though because it gives us an idea of the amount of cooperativity. Any value for  $n$  greater than 1 means that there is some cooperativity. Usually, you can guess that the number of binding sites (or the “valency”) is at *least*  $\text{ceil}(n)$  (try this in MATLAB<sup>©</sup>) and could even be higher. A value of  $n=3.2$  in this case indicates that there are at least 4 binding sites and they have a high, though not perfect, cooperativity between them.

$K_d = K_M^n$  for the fully cooperative binding case

Example problem for 2 binding sites (could be generalized to n).



$$K_1 = \frac{[P][L]}{[PL]} \quad K_2 = \frac{[PL][L]}{[PL_2]} \quad \text{and the macroscopic } K_d = \frac{[P][L]}{[PL_2]}$$

$v = \frac{\text{mol ligand bound}}{\text{mol total protein}}$

$$v = \frac{[PL] + 2[PL_2]}{[P] + [PL] + [PL_2]}$$

rearrange

$$[PL] = \frac{[P][L]}{K_1} \quad \text{and} \quad [PL_2] = \frac{[PL][L]}{K_2}$$

plug in

$$[PL_2] = \frac{[P][L]^2}{K_1 K_2}$$

$$v = \frac{\frac{[P][L]}{K_1} + \frac{2[P][L]^2}{K_1 K_2}}{[P] + \frac{[P][L]}{K_1} + \frac{[P][L]^2}{K_1 K_2}}$$

$$v = \frac{\frac{[L]}{K_1} + \frac{2[L]^2}{K_1 K_2}}{1 + \frac{[L]}{K_1} + \frac{[L]^2}{K_1 K_2}}$$

In the case of high cooperativity,  $K_2 \ll K_1$ , so  $[PL_2] \gg [PL]$  and

$$\frac{[P][L]^2}{K_1 K_2} \gg \frac{[P][L]}{K_1} \quad \text{and} \quad \frac{[L]^2}{K_1 K_2} \gg \frac{[L]}{K_1}$$

$$\text{so } v = \frac{\frac{2[L]^2}{K_1 K_2}}{1 + \frac{[L]^2}{K_1 K_2}} = \frac{2[L]^2}{K_1 K_2 + [L]^2}$$

and if  $y =$  fraction of protein binding sites with ligand bound,  
 ~~$v = 2y$~~

Generalizing for n,  $v = ny$

$$v = \frac{n[L]^n}{K_1 K_2 \dots K_n + [L]^n} \quad \text{and if } K_M = K_1 = K_2 = \dots = K_n \quad v = \frac{n[L]^n}{K_M^n + [L]^n} = \frac{n[L]^n}{K_d + [L]^n}$$

as shown in class

```
1  %problem 5
2  global Kd;
3
4  load 'pset4_data';
5  L = pset4_data(:,1);
6  Ydata = pset4_data(:,2);
7
8  Kd = 1e-4; %M = microscopic binding constant
9
10 n_guess = 4;
11 [n_fit, R, J] = nlinfit(L,Ydata,@Y_independent,n_guess);
12 n_fit
13 for i=1:length(L)
14     Yfit1(i) = (n_fit.*L(i))./(L(i) + Kd);
15 end
16
17 figure(1);
18 plot(L,Ydata,'r*');
19 hold on
20 plot(L,Yfit1,'b');
21 legend('Y data','Y fit 1');
22 xlabel('[L] M');
23 ylabel('Y');
24 title('Initial data and fit for Y versus L, independent binding assumption');
25
26 figure(2);
27 plot(L,R,'b*');
28 xlabel('[L]');
29 ylabel('Residuals');
30 title('Residuals from fit for Y versus L, independent binding assumption');
31
32 n_guess = 4;
33 [n_fit, R, J] = nlinfit(L,Ydata,@Y_cooperative,n_guess);
34 n_fit
35 for i=1:length(L)
36     Yfit2(i) = (L(i).^n_fit)./(L(i).^n_fit+Kd.^n_fit);
37 end
38
39 figure(3);
40 plot(L,Ydata,'r*');
41 hold on
42 plot(L,Yfit2,'b');
43 legend('Y data','Y fit 2');
44 xlabel('[L] M');
45 ylabel('Y');
46 title('Initial data and fit for Y versus L, cooperative binding assumption');
47
48 figure(4);
49 plot(L,R,'b*');
```

```
50 xlabel('[L]');  
51 ylabel('Residuals');  
52 title('Residuals from fit for Y versus L, cooperative binding assumption');  
53  
54
```

```
1 %fractional saturation Y for independent binding assumption
2 function y = Y_independent(n_tofit, L)
3 global Kd;
4 y = (n_tofit.*L)./(L + Kd);
```

```
1 %fractional saturation Y for cooperative binding assumption
2 function y = Y_cooperative(n_tofit, L)
3 global Kd;
4 y = (L.^n_tofit)./(L.^n_tofit+Kd.^n_tofit);
5
6
7
8
9
```