Binding Constant Determination

References:

Derivations of Binding Equations:

1. Uzunova, V. D. Thermodynamic and Fluorescence-based Investigations of Host-Dye Reporter PAirs for Sensing of Biological Analytes. Jacobs University, 2011.

Papers using derived equations (or derivatives):

2. Hennig, A.; Hoffmann, A.; Borcherding, H.; Thiele, T.; Schedler, U.; Resch-Genger, U., Quantification of surface functional groups on polymer microspheres by supramolecular host-guest interactions. Chem. Commun. (Cambridge, U. K.) 2011, 47 (27), 7842-7844.
3. Hennig, A.; Bakirci, H.; Nau, W. M., Label-free continuous enzyme assays with macrocycle-fluorescent dye complexes. Nat Meth 2007, 4 (8), 629-632.

Paper detailing how to program algorithms for fitting binding data:

4. Hargrove, A. E.; Zhong, Z.; Sessler, J. L.; Anslyn, E. V., Algorithms for the determination of binding constants and enantiomeric excess in complex host : guest equilibria using optical measurements. New J. Chem. 2010, 34 (2), 348-354.

1:1 Binding

Given a Host (H) and a Guest (G) forming an inclusion complex (HG) with an equilibrium constant K_G, and total concentrations H_0 and G_0:

$$K_G = \frac{[HG]}{[H][G]} \qquad (1)$$

$$[G] = [G]_0 - [HG] \qquad (2)$$

$$[H] = [H]_0 - [HG] \qquad (3)$$

$$K_G = \frac{[HG]}{([G]_0 - [HG])([H]_0 - [HG])} \qquad (4)$$

$$[HG]^2 + [HG]\left(-[G]_0 - [H]_0 - \frac{1}{K_G}\right) + [G]_0[H]_0 = 0 \qquad (5)$$

which fits the form : $x^2 + bx + c = 0$, and gives as a solution:

$$x = \frac{-b - \sqrt{(b)^2 - 4c}}{2}, b = -[G]_0 - [H]_0 - \frac{1}{K_G}, c = [G]_0[H]_0, \text{ to give:}$$

$$[HG] = \frac{\left([G]_0 + [H]_0 + \frac{1}{K_G}\right) - \sqrt{\left([G]_0 + [H]_0 + \frac{1}{K_G}\right)^2 - 4[G]_0[H]_0}}{2} \qquad (6)$$

$$[G] = [G]_0 - [HG], \Rightarrow [G] = \frac{[G]_0 - [H]_0 - \frac{1}{K_G} + \sqrt{\left([G]_0 + [H]_0 + \frac{1}{K_G}\right)^2 - 4[G]_0[H]_0}}{2} \qquad (7)$$

Given a partially quenched guest with intensity I_G, and the fluorescent host-guest complex with intensity I_{GH}, the overall fluorescence intensity of the system can be written as a linear combination of the mole fraction of each component with their intensity, which can in turn be related to the concentration of guest in solution:

$$\Delta I = Fl = x_G I_G + x_{GH} I_{GH} = \frac{[G]}{[G]_0} I_G + \frac{[HG]}{[G]_0} I_{GH} = I_{GH} + (I_G - I_{GH}) \frac{[G]}{[G]_0} \quad (8)$$

Combining (7) and (8) yields the final equation to fit the intensity change to find K_G:

$$Fl = I_{GH} + (I_G - I_{GH}) \frac{[G]_0 - [H]_0 - \frac{1}{K_G} + \sqrt{\left([G]_0 + [H]_0 + \frac{1}{K_G}\right)^2 - 4[G]_0[H]_0}}{2[G]_0} \quad (9)$$

To fit the data, the values for I_{GH} and I_G are determined as the maximum and minimum values from the plot (completely complexed HG and complete free G, respectively) assuming the host has no intrinsic fluorescence. The total fluorescence of the system is then plotted vs. concentration of total host, $[H]_0$, given a starting guest concentration $[G]_0$. The best fit is obtained with a slight allowance for variability of the I_G, I_{GH} parameters.

Competitive Binding:

Given a Host (H), an indicating Guest (G), and a competitive guest (C) forming two competing inclusion complexes (HG & HC) with equilibrium constants K_G & K_C, and total concentrations H_0, G_0 & C_0:

$$K_G = \frac{[HG]}{[H][G]}, K_C = \frac{[HC]}{[H][C]} \quad (1)$$

$$[G] = [G]_0 - [HG] \quad (2)$$

$$[C] = [C]_0 - [HC] \quad (3)$$

$$[H] = [H]_0 - [HG] - [HC] \quad (4)$$

$$[HG] = \frac{K_G[G]_0[H]}{1+K_G[H]}, [HC] = \frac{K_C[C]_0[H]}{1+K_C[H]} \quad (5)$$

$$[H] = [H]_0 - \frac{K_G[G]_0[H]}{1+K_G[H]} - \frac{K_C[C]_0[H]}{1+K_C[H]} \quad (6)$$

$$K_G = \frac{[HG]}{([G]_0 - [HG])([H]_0 - [HG])} \quad (4)$$

$$K_C K_G [H]^3 + \left(K_C + K_G + K_C K_G([G]_0 + [C]_0 - [H]_0)\right)[H]^2 + (K_C([C]_0 - [H]_0) + K_G([G]_0 - [H]_0) + 1)[H] - [H]_0 = 0 \quad (5)$$

which fits the form : $ax^3 + bx^2 + cx + d = 0$, with constants:

$$a = K_C K_G \quad (6)$$

$$b = K_C + K_G + K_C K_G([G]_0 + [C]_0 - [H]_0) \quad (7)$$

$$c = K_C([C]_0 - [H]_0) + K_G([G]_0 - [H]_0) + 1 \quad (8)$$

$$d = -[H]_0 \quad (9)$$

Unlike the quadratic equation that results from 1:1 binding, this cubic equation cannot be solved analytically, which leaves a computational solution. The Newton-Rhapson method solves a trinomial by using an iterative method, wherein the ratio of the polynomial to its first derivative should go to zero as the correct value of x is approached.

$$x_{n+1} = x_n - \frac{f(x_n)}{f\prime(x_n)} \qquad (10)$$

The solution can be scripted to solve for a value of x, and plug it into a fitting equation, and iterate the process based on chi-squared value.

Given a partially quenched guest with intensity I_G, and the fluorescent host-guest complex with intensity I_{GH}, the overall fluorescence intensity of the system can be written as a linear combination of the mole fraction of each component with their intensity, which can in turn be related to the concentration of guest in solution:

$$\Delta I = Fl = x_G I_G + x_{GH} I_{GH} = \frac{[G]}{[G]_0} I_G + \frac{[HG]}{[G]_0} I_{GH} = I_G + (I_{GH} - I_G)\frac{K_G[H]}{1+K_G[H]} \qquad (11)$$

Given the intensity of the mixture, the polynomial (5) can be solved using the Newton method, and iteratively advanced by fit of the resulting value for [H] to the data using (11). The following code (labscript) will work with OriginPro 9.1 to script the process, iterating the solution to the trinomial until the step function has a value of less than 1E-15 before plugging it into (11):

```
a=Kg*Kc;
b=Kg+Kc+Kg*Kc*G0-Kg*Kc*H0+Kg*Kc*Cs;
c=Kg*G0-Kg*H0-Kc*H0+Kc*Cs+1;
d=-H0;

for (H=H0, step=1; abs(step)>1E-15; H=H-step){
    step=(a*H*H*H+b*H*H+c*H+d)/(3*a*H*H+2*b*H+c);};
f=Ig+((Igh-Ig)*(kg*H)/(1+kg*H))
```

With parameters the independent variable Cs (total concentration of competitive guest), and the dependent variable f, and parameters Kc, Ig, Igh, Kg, H0 and G0, all but the latter two of which should be lightly varied during fitting for the best fit to the curve. Care should be taken not to over-iterate the fit, as in the case of a moderate fit, the parameters can over-extend trying to get the best fit curve.

To script in OriginC, the same parameters & variables can be defined. The script then for the equation follows the form:

```
double a; a=Kg*Kc;
double b; b=Kg+Kc+Kg*Kc*G0-Kg*Kc*H0+Kg*Kc*Cs;
double c; c=Kg*G0-Kg*H0-Kc*H0+Kc*Cs+1;
double d; d=-H0;
```

```
double step; double H; for (H=H0, step=1; abs(step)>1E-15; H=H-step){
    step=(a*H*H*H+b*H*H+c*H+d)/(3*a*H*H+2*b*H+c);}
f=Ig+((Igh-Ig)*(Kg*H)/(1+Kg*H))
```

Where the lack of the semi-colon after the second-to-last line is crucial for proper iteration in C syntax. Variables are defined as “double variable” the line before introduction (for all of the intermediate values not initially defined) where double allows a double precision floating-point number as opposed to an integer. Floating point-numbers can hold non integer values by allowing the decimal point to be variable (floating) in the space allotted to store the number.

Best fits are obtained allowing a +/- 10% variance in Igh and Ig.