# Algorithms for the determination of binding constants and enantiomeric excess in complex host : guest equilibria using optical measurements†

**Amanda E. Hargrove,[a] Zhenlin Zhong,[b] Jonathan L. Sessler*[a] and Eric V. Anslyn*[a]**

*Received (in Gainesville, FL, USA) 19th September 2009, Accepted 26th October 2009*
*First published as an Advance Article on the web 12th January 2010*
**DOI: 10.1039/b9nj00498j**



The determination of binding constants is central to many areas of research, supramolecular chemistry in particular. Traditional nonlinear regression analysis, however, cannot be applied to complex systems unless certain assumptions are undertaken, which often limits the reliability of such calculations. Our group has developed an iterative method using commercial software that allows for the rigorous determination of binding constants in a variety of systems, including 1 : 2 complexes, indicator displacement assays, and enantioselective indicator displacement assays. The improved accuracy of the values obtained in the latter case, in turn, allows for a more precise determination of *ee* in competitive equilibria.

## Introduction

The study of host–guest chemistry is a large and continually growing research field with applications in nearly every chemical discipline.[1] Host–guest complexes are often characterized in part by determining their thermodynamic properties, in particular the binding constant of the guest to the host. The determination of accurate binding constants can be critical to the applications of many such systems, which range from sensing to drug discovery and development. Unfortunately, current analyses often depend upon the assumption that the concentration of free guest in solution is mathematically equivalent to the total concentration of guest, thus allowing for the use of simple graphing methods, such as Benesi-Hildebrand plots and Scatchard plots.[2] However, this approximation is never fully valid and is particularly inappropriate in cases where the host–guest interaction is strong. While binding constants can be determined numerically using advanced programs such as HYPERQUAD,[3] these programs are often expensive and awkward for an untrained user. As described here, the iterative methods developed by our group allow for the rigorous application of binding equations by employing the nonlinear regression module present in Origin,[4] an inexpensive software package found in many laboratories, combined with a user-specified target function. While several researchers have developed unique in-house methods to determine binding constants for systems of 1 : 1 binding stoichiometry, such as the method described in detail by Hirose,[5] our method is further applicable to non-trivial systems of equilibria, including those characterized by a 1 : 2 binding stoichiometry or associated with common indicator displacement assays (IDAs).

Optical spectroscopy is one of the preferred methods for measuring host : guest thermodynamics due to its ease of use, straightforward data interpretation, and the requirement for relatively inexpensive instrumentation. For systems that obey the Beer–Lambert law (“Beer’s law”) and for which a 1 : 1 host : guest binding stoichiometry pertains, the mathematical isotherm used in the determination of binding constants ($K$) using optical spectroscopy can be solved in closed form, and many research groups employ nonlinear curve-fitting of the isotherm using commercial data analysis software programs, such as Origin. For more complex binding equilibria, such as those characterized by a 1 : 2 binding stoichiometry or involved in the IDAs regularly used by our research group,[6] the mathematical constructs needed to fit the isotherms contain a polynomial equation that can be readily solved using iterative methods that avoid the previously mentioned assumption. For example, in the equations used for fitting an indicator displacement isotherm (eqn (39) and (41)), the concentration of free host is related to the binding constants ($K_G$, $K_I$) through cubic eqn (39). In addition, the use of empirically determined values (in this case $K_I$, $\varepsilon_I$, $\varepsilon_{HI}$) to calculate other unknown values ($K_G$) can lead to decreased accuracy in the fitting.[7] In order to address this problem and avoid as many assumptions as possible, we have developed a method that combines an in-house written subprogram (“script”) with the nonlinear curve fitting operation of Origin. This method allows us to determine iteratively the parameter values in the relevant isotherms that best fit the experimental data. This approach has proved invaluable to our group in the determination of binding constants from indicator displacement assays[6] and has also facilitated calculations of enantiomeric excess using enantioselective indicator displacement assays (eIDAs).[7–9] After the script has been input into Origin,

[a] *Department of Chemistry and Biochemistry, The University of Texas at Austin, 1 University Station A5300, Austin, TX, 78712, USA. E-mail: anslyn@austin.utexas.edu, sessler@mail.utexas.edu; Fax: +1 512 471 7791; Tel: +1 512 471 0068*
[b] *Key Laboratory of Biomedical Polymers of Ministry of Education, Department of Chemistry, Wuhan University, Wuhan 430072, P. R. China. E-mail: zlzhong@whu.edu.cn; Fax: +86 27 6875 4509; Tel: +86 27 6875 4061*

† Electronic supplementary information (ESI) available: 1 : 2 Host : Guest Binding when Host is Transparent. See DOI: 10.1039/b9nj00498j

 

the determination of binding constants from titration data is as simple as importing the signal and concentration values from a standard spreadsheet. In hopes that such methods will be useful to others in the field, we herein describe in detail this script and the associated methods used by our group to determine binding constants for 1 : 1 and 1 : 2 host : guest binding stoichiometries and IDAs, as well as to determine enantiomeric excess *via* eIDAs.

## Results and discussion

The goal in each binding scenario is to derive an equation that relates the measured signal (absorbance or fluorescence) to the total concentration of the host and guest through the desired values ($K$ or *ee*). These equations are derived through a manipulation of equilibrium constant equations, mass balance equations, and signal-to-concentration relationships, such as Beer's law in the case of absorption. Unfortunately, the concentration of free host and guest in a given solution cannot be determined using known values. As a result, a rigorous application of the binding equations to the data more often than not requires the solution to two or more related equations. As such, they benefit from the use of iterative methods both to develop the underlying binding equations and to obtain the best fits to the data. Appreciating this need, we have incorporated Newton's iterative method[10] into a set of expressions and associated scripts.

We will begin our discussion with the simplest system, namely systems characterized by a 1 : 1 host : guest binding stoichiometry. We will then build from this foundation in the treatment of situations with greater complexity. While we appreciate that a number of alternative approaches may be used to treat 1 : 1 binding processes, we focus on a description that most closely mirrors the methods we use to treat more complex situations.

### 1 : 1 Binding

We begin each derivation by defining the solution equilibria and mass balance equations. In a simple host–guest binding equilibrium, assuming both the host (H) and guest (G) have only one binding site, the equilibrium equation is represented by eqn (1).

$$\mathrm{H} + \mathrm{G} \rightleftharpoons \mathrm{HG} \tag{1}$$

The desired binding constant ($K$) is expressed in eqn (2).

$$K = \frac{[\mathrm{HG}]}{[\mathrm{H}][\mathrm{G}]} \tag{2}$$

Assigning the total concentrations of H and G as $[\mathrm{H}]_t$ and $[\mathrm{G}]_t$, respectively, gives mass balance eqn (3) and (4).

$$[\mathrm{H}]_t = [\mathrm{H}] + [\mathrm{HG}] \tag{3}$$

$$[\mathrm{G}]_t = [\mathrm{G}] + [\mathrm{HG}] \tag{4}$$

Using eqn (2)–(4) as a starting point, the first step is to derive an equation based on only one unknown concentration ([G], [H], or [HG]) to which all other concentrations are related. We arbitrarily chose [G] for this example and thus begin with the modification of eqn (4). This modification first requires the definition of [HG] in terms of [G]. Eqn (2) is rearranged to define [HG] and then used to substitute for [HG] in eqn (3). Solving for [H] yields eqn (5).

$$[\mathrm{H}] = \frac{[\mathrm{H}]_t}{1 + K[\mathrm{G}]} \tag{5}$$

Combining rearranged eqn (2) and (5) then gives eqn (6).

$$[\mathrm{HG}] = \frac{K[\mathrm{G}][\mathrm{H}]_t}{1 + K[\mathrm{G}]} \tag{6}$$

Substituting eqn (6) into eqn (4) yields a quadratic equation, which is rearranged to give eqn (7).

$$K[\mathrm{G}]^2 + (1 - K[\mathrm{G}]_t + K[\mathrm{H}]_t)[\mathrm{G}] - [\mathrm{G}]_t = 0 \tag{7}$$

The real root of eqn (7) is expressed in eqn (8), which defines [G] based on $K$ and experimentally determined values ($[\mathrm{H}]_t$ and $[\mathrm{G}]_t$).

$$[\mathrm{G}] = \frac{-(1 - K[\mathrm{G}]_t + K[\mathrm{H}]_t) + \sqrt{(1 - K[\mathrm{G}]_t + K[\mathrm{H}]_t)^2 + 4K[\mathrm{G}]_t}}{2K} \tag{8}$$

We now turn to analyzing the optical response as a function of known or measurable parameters. In this case, we will discuss the derivation for an absorption experiment. Assuming that all species present follow Beer's law, the absorbance of the equilibrium solution at any given wavelength can be expressed as eqn (9), where $A$ is the absorbance at the selected wavelength; $\varepsilon_H$, $\varepsilon_G$, $\varepsilon_{HG}$, are the molar absorptivities of the host, guest, and the host : guest complex, respectively, at that wavelength; and b is the path length of the cell.

$$A = \varepsilon_H b[\mathrm{H}] + \varepsilon_G b[\mathrm{G}] + \varepsilon_{HG} b[\mathrm{HG}] \tag{9}$$

In the case that the host is transparent at the selected wavelength, $\varepsilon_H = 0$, then eqn (9) can be simplified to eqn (10). Of course, the definition of which species is the host and which is the guest is arbitrary.

$$A = \varepsilon_G b[\mathrm{G}] + \varepsilon_{HG} b[\mathrm{HG}] \tag{10}$$

Under these circumstances, the binding study is usually performed such that the total concentration of the guest ($[\mathrm{G}]_t$) is kept constant while the total concentration of the host ($[\mathrm{H}]_t$) is increased gradually. A titration curve is obtained by plotting the absorbance ($A$) at a certain wavelength against $[\mathrm{H}]_t$. Substituting eqn (6) and (8) into eqn (10) gives the final 1 : 1 binding isotherm (eqn (11)).

$$A = \left\{ \varepsilon_G b + \frac{\varepsilon_{HG} b K[\mathrm{H}]_t}{1 + 0.5\left\{ -(1 - K[\mathrm{G}]_t + K[\mathrm{H}]_t) + \sqrt{(1 - K[\mathrm{G}]_t + K[\mathrm{H}]_t)^2 + 4K[\mathrm{G}]_t} \right\}} \right\} \times \frac{-(1 - K[\mathrm{G}]_t + K[\mathrm{H}]_t) + \sqrt{(1 - K[\mathrm{G}]_t + K[\mathrm{H}]_t)^2 + 4K[\mathrm{G}]_t}}{2K} \tag{11}$$

In eqn (11), the path length (b), the molar absorptivity of the guest ($\varepsilon_G$), and the guest total concentration ($[\mathrm{G}]_t$) are known values. The binding constant ($K$) and molar absorptivity of the

complex ($\varepsilon_{HG}$) can be determined using the titration data and nonlinear regression analysis. For example, eqn (11) can be used as the input for a user-defined function in the nonlinear fitting function of Origin, which will then determine the values of $K$ and $\varepsilon_{HG}$ from the best-fit curve. In our analysis, we treat parameter $\varepsilon_G$ as slightly adjustable, since molar absorptivities are inherently dependent on a variety of experimental conditions (temperature, background absorbance, *etc.*) and cannot always be measured with an accuracy that translates from experiment to experiment. Varying this parameter, and related values in later analyses, greatly improves the accuracy of these iterative methods.

It should be noted that the derivations described here are often based on experiments in which certain species in the solution are transparent at the studied wavelength, such as the host in the case of a 1 : 1 binding stoichiometry. While such experimental conditions greatly simplify the resulting algebra, they are not necessary for the application of our method. Because we define each species concentration relative to a single unknown concentration, the omitted values can be easily substituted into the original signal-*vs.*-concentration equation and the corresponding expression solved in a manner analogous to what is described. For example, in the case of a 1 : 1 binding stoichiometry, eqn (5), (6) and (8), which relate each species to [G], would be substituted into eqn (9), which relates the absorbance to the concentration of each species. This substitution would result in a slightly longer final binding equation that can be input into the nonlinear fitting function of Origin in the same manner as described for eqn (11). This approach is applicable in all of the binding situations described here.

### 1 : 2 Host : Guest binding

The binding of two guests to a single host molecule greatly complicates the thermodynamic analysis as two binding constants ($K_1$ and $K_2$) must be determined. Whereas 1 : 1 binding derivations lead to a quadratic equation (eqn (7)), 1 : 2 binding derivations lead to a cubic equation (eqn (22)). We have found that this latter equation is solved most easily using Newton's iterative optimization method. As detailed below, we do this using a subprogram (script) generated in-house with the commercially available Origin software program. This script and the relevant equations are discussed below.

The derivation begins similarly to that of 1 : 1 binding. In this situation there are two equilibrium equations (eqn (12), (13)), two binding constant expressions (eqn (14), (15)), and two mass balance equations (eqn (16), (17)).

$$H + G \rightleftharpoons HG \quad (12)$$

$$HG + G \rightleftharpoons HG_2 \quad (13)$$

$$K_1 = [HG]/([H][G]) \quad (14)$$

$$K_2 = [HG_2]/([HG][G]) \quad (15)$$

$$[H]_t = [H] + [HG] + [HG_2] \quad (16)$$

$$[G]_t = [G] + [HG] + 2[HG_2] \quad (17)$$

As before, the first goal is to derive an equation dependent on only one unknown concentration, and we again choose [G]. In a manner similar to the previous manipulation of eqn (2), (3) and (5) to give eqn (6), eqn (14)–(16) are combined to give eqn (18)–(20).

$$[H] = \frac{[H]_t}{1 + K_1[G] + K_1K_2[G]^2} \quad (18)$$

$$[HG] = \frac{K_1[G][H]_t}{1 + K_1[G] + K_1K_2[G]^2} \quad (19)$$

$$[HG_2] = \frac{K_1K_2[G]^2[H]_t}{1 + K_1[G] + K_1K_2[G]^2} \quad (20)$$

Substituting eqn (19) and (20) into eqn (17) yields eqn (21), in which the only unknown variables are [G], $K_1$, and $K_2$.

$$[G_t] = [G] + \frac{K_1[G] + 2K_1K_2[G]^2}{1 + K_1[G] + K_1K_2[G]^2}[H]_t \quad (21)$$

Rearranging eqn (21) results in a cubic equation for [G], which for convenience we express as eqn (22).

$$A[G]^3 + B[G]^3 + C[G] + D = 0 \quad (22)$$

where

$$A = K_1K_2$$

$$B = K_1 + 2K_1K_2[H]_t - K_1K_2[G]_t$$

$$C = 1 + K_1[H]_t - K_1[G]_t$$

$$D = -[G]_t$$

With the [G] dependent equation in hand, we again derive a relationship to the optical data. Accordingly, eqn (23) is derived from Beer's law.

$$A = \varepsilon_H b[H] + \varepsilon_G b[G] + \varepsilon_{HG} b[HG] + \varepsilon_{HG_2} b[HG_2] \quad (23)$$

As an example of derivation with the opposite absorbance pattern, we will now assume the host has an absorbance and the guest is transparent at the observed wavelength. In this case, a titration is usually performed by keeping $[H]_t$ constant and incrementally increasing $[G]_t$. As a result, eqn (23) can be simplified to eqn (24). By substituting eqn (18)–(20) into eqn (24), eqn (25) is finally obtained, in which the absorbance is dependent on only one concentration.

$$A = \varepsilon_H b[H] + \varepsilon_{HG} b[HG] + \varepsilon_{HG_2} b[HG_2] \quad (24)$$

$$\boxed{A = \frac{\varepsilon_H b + \varepsilon_{HG} bK_1[G] + \varepsilon_{HG_2} bK_1K_2[G]^2}{1 + K_1[G] + K_1K_2[G]^2}[H]_t} \quad (25)$$

An experimentally generated isotherm is obtained by plotting the absorbance at a specific wavelength against the total guest concentration ($A$ *vs.* $[G]_t$). Though the value of [G] in eqn (25) varies depending on $[G]_t$ (as described by eqn (22)), the solution can be readily approached through an iterative data fitting procedure.

Specifically, we first apply Newton's method to iteratively determine numerical values for [G] (eqn (26), $x$ = [G]). We herein summarize Newton's method to the extent we use it in

 

the determination of binding constants and related variables. In brief, the true value ($x$), the root of the equation, is approximated using an iterative method based on eqn (26). In the current case, $f(x)$ is the left side of eqn (22), and $f'(x)$ is its derivative. Since the root ($x$) is represented by the $x$-intercept ($y = 0$), $f(x_n)$ approaches zero as the estimated values ($x_n$) become closer to the true value. In this equation, an initial guess ($n = 0$) is used as the input to determine the first iteration value ($x_1$). The first iteration value ($x_1$) is then used to determine $x_2$ and so on until the second term of eqn (26) becomes very close to zero, at which point $x_{n+1}$ equals $x_n$, and the mathematically approximate value of $x$ is found. This second term is often referred to as the "step" by which iterative values approach the true value. While such an iterative method can be very powerful, it should be noted that its success depends upon the judicious choice of the initial guess for each parameter. The initial value must be reasonably close to the true values in order for the step function to approach it reliably.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{26}$$

With all of the necessary equations in hand, a script can be written in the nonlinear curve fitter of Origin to numerically solve eqn (22) for [G] while fitting the data using eqn (25). This script, shown below, fits a set of titration data in which absorbance ($y$ in the script) is recorded as a function of $[G]_t$ ($x$ in the script).[11] $[G]_t$ is used as the initial guess for [G] (*i.e.*, $x_0 = [G]_t$):

```
A=K1*K2;
B=K1+2*K2*K1*Ht-K1*K2*x;
C=1+K1*Ht-K1*x;
D=-x;

for (G=x, step=1; abs(step)>1e-15; G=G-step){
   step=(a*G*G*G+b*G*G+c*G+d)/(3*a*G*G+2*b*G+c);};

y=Ht*(E0+E1*K1*G+E2*K1*K2*G*G)/(1+K1*G+K1*K2*G*G)
```

where K1, K2, E0, E1, E2, Ht, x, and y represent $K_1$, $K_2$, $\varepsilon_H b$, $\varepsilon_{HG} b$, $\varepsilon_{HG2} b$, $[H]_t$, $[G]_t$, and absorbance ($A$), respectively. Parameters E0 and Ht ($\varepsilon_H b$ and $[H]_t$) are known, though $\varepsilon_H$ was again treated as slightly adjustable. Values for x and y ($[G]_t$ and $A$) are read from a standard data file by the program. This script sets up an IF/THEN statement in which the absolute value of the step function is evaluated. As discussed in the context of eqn (26), the step is defined (step =) as $f(x_n)/f'(x_n)$ in which $f(x)$ is the left side of eqn (22). ***If*** this value is significantly greater than zero (abs(step) $> 1 \times 10^{-15}$), ***then*** another iteration is performed (G = G-step). Once the step value is effectively zero for each data point, the resulting solutions for [G] are used to determine the best fit curve with eqn (25), represented in the last line of the script. If the least squares regression analysis yields a sufficiently low value, the process is complete. If not, improved values for the unknown parameters (K1, K2, E1, and E2) are used in another cycle, beginning again with the iterative determination of new [G] values. Given a set of estimated initial values for K1, K2, E1, and E2, the program iteratively determines the parameter values that best fit the experimental data. A similar situation is observed when H is transparent and is described in the ESI.†

### Indicator displacement

In an indicator displacement assay, a host is first allowed to complex with a dye (indicator, I) that displays a change in absorbance or fluorescence upon binding. The addition of guest then displaces the indicator, leading to a reversal in the absorption or fluorescence spectrum (eqn (27)). The degree of displacement is used to calculate the affinity of the guest to the host ($K_G$). This assay does not require the host or guest to undergo changes in its optical properties in order to measure binding. In the situation presented here, the host : indicator complex as well as the host : guest complex display 1 : 1 binding stoichiometries.

$$HI + G \rightleftharpoons HG + I \tag{27}$$

Following the previous protocols, dye displacement systems involve two independent equilibria (eqn (28), (29)), two binding constant expressions (eqn (30) and (31)), and three mass balance equations (eqn (31)–(33)).

$$H + G \rightleftharpoons HG \tag{28}$$

$$H + I \rightleftharpoons HI \tag{29}$$

$$K_G = [HG]/([H][G]) \tag{30}$$

$$K_I = [HI]/([H][I]) \tag{31}$$

$$[H]_t = [H] + [HG] + [HI] \tag{32}$$

$$[G]_t = [G] + [HG] \tag{33}$$

$$[I]_t = [I] + [HI] \tag{34}$$

We again use these equations as a starting point to derive an equation involving only one unknown concentration, and in this case we chose [H]. We therefore focus on eqn (32) and seek to define all other concentrations in terms of [H]. Accordingly, combining eqn (30)–(34) yields eqn (35)–(37).

$$[HG] = \frac{K_G[H]}{1 + K_G[H]}[G]_t \tag{35}$$

$$[HI] = \frac{K_I[H]}{1 + K_I[H]}[I]_t \tag{36}$$

$$[I] = \frac{I_t}{1 + K_I[H]} \tag{37}$$

Substituting eqn (35) and (36) into eqn (32) in turn leads to eqn (38).

$$[H]_t = [H] + \frac{K_G[H]}{1 + K_G[H]}[G]_t + \frac{K_I[H]}{1 + K_I[H]}[I]_t \tag{38}$$

Eqn (38) is a cubic equation for [H], which can be rearranged to its polynomial form to give eqn (39).

$$A[H]^3 + B[H]^2 + C[H] + D = 0 \tag{39}$$

where

$$A = K_I K_G$$

$$B = K_I + K_G + K_I K_G[I]_t + K_I K_G[G]_t - K_I K_G[H]_t$$

$$C = 1 + K_I[I]_t + K_G[G]_t - K_I[H]_t - K_G[H]_t$$

$$D = -[H]_t$$

In a traditional indicator displacement titration, the total concentrations of the host and indicator are kept constant, and the total concentration of the guest is incrementally increased. An experimentally generated isotherm is obtained by plotting the absorbance at a specific wavelength against total guest concentration (*A vs.* $[G]_t$). Therefore, the next step in the derivation is to relate the optical signal, absorbance, to [H] using Beer's law. Because it simplifies the mathematics, the IDA analysis is usually performed using host and guest molecules that are transparent at the observed wavelength. To the extent such an assumption is valid, the absorbance can be related solely to the concentrations of the free indicator ([I]) and the complexed indicator ([HI]) as in eqn (40). Substitution of eqn (36) and (37) into eqn (40) yields final eqn (41).

$$A = \varepsilon_I b[I] + \varepsilon_{HI} b[HI] \quad (40)$$

$$A = \frac{[I]_t}{1 + K_I[H]}(\varepsilon_I b + \varepsilon_{HI} b K_I[H]) \quad (41)$$

As [H] varies at different values of $[G]_t$, Newton's method is again applied. Analogous to the script for 1 : 2 binding, the user-defined script given below is based on eqn (39) and (41) where $[H]_t$ is used as the initial guess for [H]:

```
A=Ki*Kg;
B=Ki+Kg+Ki*Kg*It+ Ki*Kg*x- Ki*Kg*Ht;
C=1+Ki*It + Kg*x-(Ki+Kg)*Ht;
D=-Ht;

for (H=Ht, 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);};

y=It*(Ei+Ehi*Ki*H)/(1+Ki*H)
```

where Ki, Kg, Ei, Ehi, Ht, It, *x*, and *y* represent $K_I$, $K_G$, $\varepsilon_I b$, $\varepsilon_{HI} b$, $[H]_t$, $[I]_t$, $[G]_t$, and absorbance (*A*), respectively. Parameters $K_I$, $\varepsilon_I$ and $\varepsilon_{HI}$ are determined by the titration and 1 : 1 curve fitting of the host and indicator alone prior to this analysis. Each of these values, however, is treated as adjustable in accordance with previous discussions. $[H]_t$ and $[I]_t$ are treated as constants, and the variables $[G]_t$ (*x*) and absorbance (*y*) are read from the data sheet by the program. Given an estimated initial value for $K_G$, the program can iteratively find the value of $K_G$ that best fits the experimental data.

### *ee* Measurement by dye-displacement

By using a chiral host (H*), enantiomeric excess (*ee*) of a guest can also be determined using an IDA. In this situation, the complexation of a chiral host with two enantiomers of the guest will lead to two diastereomers. The differing stabilities of these diastereomers in turn produce different degrees of indicator displacement (eqn (44) and (46)). The signal at different *ee*'s thus depends upon the two equilibrium constants $K_R$ and $K_S$ as defined in eqn (49) and (51). As discussed in previous work,[7] though these values could be determined independently, the most accurate *ee* determination is achieved when $K_R$ and $K_S$ are calculated using iterative methods based on titration data at varying *ee* values. We begin, as before, by defining the equilibrium, binding constant, and mass balance expressions for this system (eqn (42)–(55)).

$$H^* + I \rightleftharpoons H^*I \quad (42)$$

$$H^* + G_R \rightleftharpoons H^*G_R \quad (43)$$

$$H^*I + G_R \rightleftharpoons H^*G_R + I \quad (44)$$

$$H^* + G_S \rightleftharpoons H^*G_S \quad (45)$$

$$H^*I + G_S \rightleftharpoons H^*G_S + I \quad (46)$$

$$K_I = [H^*I]/([H^*][I]) \quad (47)$$

$$K_{GR} = [H^*G_R]/([H^*][G_R]) \quad (48)$$

$$K_R = K_{GR}/K_I = [H^*G_R][I]/([H^*I][G_R]) \quad (49)$$

$$K_{GS} = [H^*G_S]/([H^*][G_S]) \quad (50)$$

$$K_S = K_{GS}/K_I = [H^*G_S][I]/([H^*I][G_S]) \quad (51)$$

$$[H^*]_t = [H^*] + [H^*I] + [H^*G_R] + [H^*G_S] \quad (52)$$

$$[G_R]_t = [G_R] + [H^*G_R] \quad (53)$$

$$[G_S]_t = [G_S] + [H^*G_S] \quad (54)$$

$$[I]_t = [I] + [H^*I] \quad (55)$$

In this derivation, we relate all concentrations to [H*I]. To begin, we combine eqn (47), (52) and (55) to give eqn (56).

$$K_I = [H^*I]/(([H^*]_t - [H^*G_R] - [H^*G_S] - [H^*I])([I]_t - [H^*I])) \quad (56)$$

We next need to define $[H^*G_R]$ and $[H^*G_S]$ in terms of [H*I]. Substituting eqn (53) and (55) into eqn (49) yields eqn (57).

$$K_R = [H^*G_R]([I]_t - [H^*I]/([H^*I]([G_R]_t - [H^*G_R])) \quad (57)$$

Eqn (57) can then be solved for $[H^*G_R]$, which results in eqn (58). A similar process can be used to derive eqn (59).

$$[H^*G_R] = K_R[H^*I][G_R]_t/([I]_t + (K_R - 1)[H^*I]) \quad (58)$$

$$[H^*G_S] = K_S[H^*I][G_S]_t/([I]_t + (K_S - 1)[H^*I]) \quad (59)$$

After substituting eqn (58) and (59) into eqn (56), the resulting equation is a 4th order polynomial (eqn (60)).

$$A[H^*I]^4 + B[H^*I]^3 + C[H^*I]^2 + D[H^*I] + E = 0 \quad (60)$$

 

where

$$A = K_I(K_R - 1)(K_S - 1)$$

$$B = K_R - 1 - [G_R]_t K_I K_R + K_S - [G_S]_t K_I K_S - K_R K_S + [G_R]_t K_I K_R K_S + [G_S]_t K_I K_R K_S + [H^*]_t K_I(K_R + K_S - 1 - K_R K_S) + [I]_t K_I(2K_R + 2K_S - 3 - K_R K_S)$$

$$C = -[I]_t[K_R - 2 - 2[G_R]_t K_I K_R + K_S - 2[G_S]_t K_I K_S + [G_R]_t K_I K_R K_S + [G_S]_t K_I K_R K_S + [I]_t K_I(K_R + K_S - 3) + [H^*]_t K_I(2K_R - K_R K_S - 3 + 2K_S)]$$

$$D = -[I]_t^2[1 + [I]_t K_I + [G_R]_t K_I K_R + [G_S]_t K_I K_S - [H^*]_t K_I(K_R + K_S - 3)]$$

$$E = [H^*]_t[I]_t^3 K_I$$

While the previous examples monitored changes in the absorption spectrum throughout a titration, any of our methods can also be applied when using fluorescence spectroscopy. We present such a scenario here as an example of an additional application of our methods. We first relate the fluorescence signal to the concentration of [H*I]. Assuming that the host and guest have negligible fluorescence at the wavelength studied, the fluorescence intensity can be described by eqn (61) ($I_0$, the intensity of the excitation source; $\varphi$, fluorescence quantum yield; ε, molar absorptivity at the excitation wavelength; $b$, path length).[12] The initial fluorescence ($F_0$) is defined in eqn (62). The corrected fluorescence ($F/F_0$) can then be derived from eqn (61) and (62), resulting in eqn (63). This equation can be further simplified to eqn (64) using mass balance eqn (55), in which the signal depends only on [H*I].

$$F = k_I b[I] + k_{HI} b[H^*I] \qquad (k \equiv 2.3\, I_0 \varphi \varepsilon) \qquad (61)$$

When $[H^*]_t = 0$,

$$F_0 = k_I b[I]_t \qquad (62)$$

$$F/F_0 = [I]/[I]_t + n[H^*I]/[I]_t \qquad (n \equiv k_{HI}/k_I) \qquad (63)$$

$$F/F_0 = 1 + (n-1)[H^*I]/[I]_t \qquad (64)$$

In this example, [H*I] of a given solution varies depending on $ee_R$ and $[G]_t$. Values for [H*I] can be iteratively determined from eqn (60) using Newton's method with $0.0001I_t$ as an initial [H*I] value. These results can then be used for the curve fitting of eqn (64), which ultimately leads to the iterative determination of $K_R$ and $K_S$ (see below). Fluorescence intensity data ($y$) is now collected in terms of $ee_R$ ($x$), not as a function of concentration as in our other examples. Because the binding constants are expressed in terms of $[G_R]_t$ and $[G_S]_t$, these concentrations must be related to $[G]_t$ according to the definition of $ee$ (eqn (65)–(66)).

$$[G_R]_t = 0.5(1 + ee_R)[G]_t \qquad (65)$$

$$[G_S]_t = 0.5(1 + ee_S)[G]_t \qquad (66)$$

Using eqn (49), (51), (60), (64), (65), and (66), the following script can be generated.

```
Grt=(0.5+x/2)*Gt; Gst=(0.5-x/2)*Gt;
Kr=Kgr/Ki; Ks=Kgs/Ki;

A=Ki*(Kr-1)(Ks-1);
B=Kr-1-Grt*Ki*Kr+Ks-Gst*Ki*Ks-Kr*Ks+Grt*Ki*Kr*Ks+
Gst*Ki*Kr*Ks+Ht*Ki*(Kr-1+Ks-Kr*Ks)+It*Ki*(-3+2*Kr+
2*Ks-Kr*Ks);
C=-It*(-2+Kr-2*Grt*Ki*Kr+Ks-2*Gst*Ki*Ks+Grt*Ki*Kr*Ks
+Gst*Ki*Kr*Ks+It*Ki*(-3+Kr+Ks)+Ht*Ki*(-3-Kr*(-2+Ks)
+2*Ks));
D=-It*It*(1+It*Ki+Grt*Ki*Kr+Gst*Ki*Ks-Ht*Ki*(-3+Kr+Ks));
E=Ht*It*It*It*Ki;

for (HI=It*0.0001, step=1; abs(step)>1e-14; HI=HI-step){
step=(a*HI^4+b*HI*HI*HI+c*HI*HI+d*HI+e)/
 (4*a*HI*HI*HI+3*b*HI*HI+2*c*HI+d);};

y=(It-HI)/It+n*HI/It
```

where Grt, Gst, Gt, Kr, Kgr, Ks, Kgs, Ki, Ht, It, and HI represent $[G_R]_t$, $[G_S]_t$, $[G]_t$, $K_R$, $K_{GR}$, $K_S$, $K_{GS}$, $K_I$, $[H^*]_t$, $[I]_t$ and [H*I], respectively. In this example, the first two lines define variable relationships based on eqn (49), (51), (65) and (66). In a manner analogous to previous protocols, the next section of script uses Newton's method to generate values for [H*I] based on eqn (60). Finally, the last line uses the generated values of [H*I] and eqn (64) to fit the data. In this case, parameters $[H^*]_t$, $[I]_t$ and $[G]_t$ are treated as constants, and the variables $ee$ ($x$) and fluorescence intensity ($y$) are read from the data sheet by the program. Parameters $n$ and $K_I$ are determined through titration and 1 : 1 curve fitting of the host and indicator alone prior to this analysis. Unlike the case of traditional IDAs, these two experimentally calculated parameters are treated as constants in the script so that $K_R$ and $K_S$ can be allowed to vary slightly. The values of $K_R$ and $K_S$ can then be input into eqn (67), which relates the corrected fluorescence signal directly to $ee$. The derivation of eqn (67) has been described by our research group.[7]

$$[H]_t = \frac{(F/F_0 - 1)[I]_t}{n-1} + \frac{(F/F_0 - 1)}{K_I(n - F/F_0)} + \frac{K_R[G]_t(1 + ee_R)(1 - F/F_0)}{2[(F/F_0)(K_I - K_R) - (nK_I - K_R)]} + \frac{K_S[G]_t(1 + ee_R)(1 - F/F_0)}{2[(F/F_0)(K_I - K_S) - (nK_I - K_S)]} \qquad (67)$$

By inputting the derived $K$ and ε values and measured fluorescence intensity, eqn (67) can be solved using a common graphing calculator to give accurate $ee$ values.[13] As previously mentioned, while $K_R$ and $K_S$ can also be determined through displacement assays with pure solutions of either the $R$ or $S$ guest, more accurate values of $ee$ are obtained from eqn (67) when the equilibrium constants are determined through the above iterative technique.

## Conclusions

The measurement of binding constants is among the most fundamental tasks involved in many supramolecular chemistry

studies, and therefore the mathematical derivation of isotherms for modelling complex equilibria is an important part of the field. Furthermore, optical spectroscopy is one of the most commonly used experimental methods to determine binding constants. Here, we have detailed several examples of the general strategy that our group uses to model complex equilibria using experimental data obtained through optical means. The approach commences by deriving equations that relate the concentration of each species in solution to one unknown concentration. This derivation results in a polynomial whose order depends upon the complexity of the binding process and whose real root is the concentration of the single unknown species. When the polynomial is higher than second order, as in the more complex equilibria involved in the formation of a complex with 1 : 2 host : guest stoichiometry or associated with standard IDAs, Newton's method is used to solve this polynomial. Then, using an equation that relates the optical data to the selected unknown concentration, the binding constants and other unknown parameters are iterated to achieve the best fit curve. This two-step iterative process can be performed with the commercially available software Origin. Once the mathematical isotherm for a particular set of interrelated and complex equilibria has been derived, and the script has been entered into Origin, the technique allows binding constants to be determined quickly and accurately without having to rely on the assumptions, such as $[H] = [H]_t$ or $[G] = [G]_t$, that are inherent in many other common treatment methods.



## Acknowledgements

This work was supported by the Welch Foundation (grants F1151 and F1193 to E.V.A.) and the National Institutes of Health (grants GM077437 and EB00549 to E.V.A. and grant GM58907 to J.L.S.).

## References

1 *Supramolecular Chemistry*, ed. J. W. Steed and J. L. Atwood, John Wiley & Sons, Ltd., West Sussex, UK, 2nd edn, 2009.
2 K. A. Connors, *Binding Constants: The Measurements of Molecular Complex Stability*, 1987.
3 P. Gans, A. Sabatini and A. Vacca, *Talanta*, 1996, **43**, 1739–1753.
4 http://www.originlab.com/.
5 K. Hirose, *J. Inclusion Phenom. Macrocyclic Chem.*, 2001, **39**, 193–209.
6 B. T. Nguyen and E. V. Anslyn, *Coord. Chem. Rev.*, 2006, **250**, 3118–3127.
7 L. Zhu, Z. Zhong and E. V. Anslyn, *J. Am. Chem. Soc.*, 2005, **127**, 4260–4269.
8 L. Zhu and E. V. Anslyn, *J. Am. Chem. Soc.*, 2004, **126**, 3676–3677.
9 L. Zhu, S. H. Shabbir and E. V. Anslyn, *Chem.–Eur. J.*, 2007, **13**, 99–104.
10 F. S. Acton, *Numerical Methods That Work*, Harper & Row, New York, 1970.
11 The use of capital or lower case letters in scripts depends entirely on the preference of the user and does not affect the calculations. In this case, capital letters were deemed necessary for clarity.
12 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Plenum Press, New York, 1999.
13 The commercial software program Mathematic (http://www.wolfram.com) was used by our group to both derive eqn (67) and to calculate enantiomeric excess values from fluorescence measurements using this equation.

 