

Quantitative Analysis of Protein–Protein Interactions

Keith D. Wilkinson

Abstract

Numerous authors, including contributors to this volume, have described methods to detect protein–protein interactions. Many of these approaches are now accessible to the inexperienced investigator thanks to core facilities and/or affordable instrumentation. This chapter discusses some common design considerations that are necessary to obtain valid measurements, as well as the assumptions and analytical methods that are relevant to the quantitation of these interactions.

Key Words

Ligand binding; protein–protein interaction; fluorescence; binding equations; binding equilibria.

1. Introduction

In the post-genomic era, the importance of protein–protein interactions is becoming even more apparent (*1*). We are coming to recognize that most, if not all, catalytic and regulatory pathways operate as networks, with frequent and extensive input from signaling pathways, feedback, and cross-talk. Replication, transcription, translation, signal transduction, protein trafficking, and protein degradation are all accomplished by protein complexes, often temporally assembled and disassembled to accomplish vectoral processes. Often these interactions are driven by interaction of recognized domains in the constituent proteins (Chapter 1). We must identify and understand these domain interactions in order to discern the patterns and logic of cellular regulation (*2*).

2. Assumptions

There are several assumptions inherent to any analysis of a simple ligand–receptor interaction (<http://www.panvera.com/tech/fpguide/FP7.pdf>).

From: *Methods in Molecular Biology*, vol. 261: *Protein–Protein Interactions: Methods and Protocols*
Edited by: H. Fu © Humana Press Inc., Totowa, NJ

1. The interactions are assumed to be reversible. In the simplest case, the association reaction is bimolecular while the dissociation reaction is unimolecular.
2. All receptor molecules are equivalent and independent.
3. The measured response is proportional to the number of occupied receptor sites.
4. The interactions are measured at equilibrium.
5. The components do not undergo any other chemical reactions and exist only in the free or bound states.

Any or all of these assumptions may prove to be unfounded in a more complex case. In fact, it is the deviation from simple behavior that is often the first indication of a more complex binding event, and each assumption should be explored to explain deviations from simple behavior. Outlined below are treatments for simple cases. A general method to obtaining binding formulas for more complex cases has been derived from statistical thermodynamic principles (3).

1.2. Binding to One Site

The receptor–ligand terminology is useful, even if artificial, in the case of protein–protein interactions. Either protein could be considered the receptor or the ligand. For the purposes of this chapter we will refer to the protein present in fixed and limiting amounts as the receptor and the component that is varied as the ligand. Thus, for one molecule of L binding to one molecule of R :



where R_f is the concentration of free receptor, L_f is the concentration of free ligand, RL is the concentration of the complex, k_1 is the association rate constant, and k_2 is the dissociation rate constant. At equilibrium,

$$\frac{[R_f][L_f]}{[RL]} = \frac{k_2}{k_1} = K_d \quad (2)$$

where K_d is the dissociation constant.

Rewriting eq. (2) in terms of total ligand $[L_t]$ and receptor concentrations $[R_t]$ and applying the conservation of mass assumption, $[L_f] = [L_t] - [RL]$ and $[R_f] = [R_t] - [RL]$, gives

$$\frac{([L_t] - [RL])([R_t] - [RL])}{[RL]} = K_d \quad (3)$$

We can rearrange eq. (3) to give the fractional saturation $[RL]/[R_t]$:

$$\frac{[RL]}{[R_t]} = \frac{[L_t] - [RL]}{K_d + [L_t] - [RL]} = \frac{[L_f]}{K_d + [L_f]} \quad (4)$$

Thus, a plot of fraction saturation $[RL]/[R_t]$ vs $[L_f]$ will give the familiar rectangular hyperbola if only one type of binding site is present (**Fig. 1A**). Alternatively, a plot of fractional saturation vs $\log[L_t]$ can be used. If free concentrations are actually measured (instead of calculated): we can use the Klotz plot (**4**), a plot of fractional saturation vs $\log[L_f]$ (**Fig. 1B**), or the Scatchard plot, a plot of ligand bound/ligand free vs ligand free (**Fig. 1C**).

1.3. Binding to Multiple Sites

It should be noted that if more than one ligand molecule binds to R , then the behavior may be more complex. For n multiple binding sites we get:

$$[RL] = [RL_1] + [RL_2] + \dots + [RL_n] = \frac{[R_t][L_f]}{K_{d1} + [L_f]} + \frac{[R_t][L_f]}{K_{d2} + [L_f]} + \dots + \frac{[R_t][L_f]}{K_{dn} + [L_f]} \quad (5)$$

where n different sites can be occupied by ligand with the corresponding binding constants.

1.3.1. Identical, Noninteracting Binding Site(s)

If all binding sites are identical and noninteracting (i.e., all bind with the same K_d), then eq. (5) reduces to

$$\frac{[L_b]}{n[R_t]} = \frac{[L_f]}{K_d + [L_f]} \quad (6)$$

where $n[R_t] = [R_f] + [L_b]$.

Note that this equation is similar to eq. (4) except for the inclusion of the stoichiometry, n . A Klotz plot of fractional saturation vs $\log[L_f]$ will be sigmoidal and symmetrical about the midpoint. The curve is nearly linear from 0.1 to $10 \times K_d$ and 99% saturation is achieved when $[L_t]$ is two orders of magnitude above K_d . A complete description of binding and accurate estimation of the plateau values requires that $[L_f]$ vary from two log units below to two log units above K_d . A steeper curve is indicative of positive cooperativity, while a flatter curve could be due to negative cooperativity or the presence of an additional binding site. The stoichiometry is calculated from the plateau value and $[R_t]$, whereas the K_d is calculated from the midpoint (5), or, more accurately, using a nonlinear least squares fit to eq. (6).

If free ligand is not measured, then we must use a plot of fractional saturation vs $\log[L_t]$, and the curve will deviate from sigmoidal by the difference between $\log[L_f]$ and $\log[L_t]$. This condition is often referred to as ligand depletion (6,7). It should be recognized, however, that it may not be possible to cover such a large range of concentrations with proteins. At the low end, we are often limited by the sensitivity of the technique and, at the high end, limited

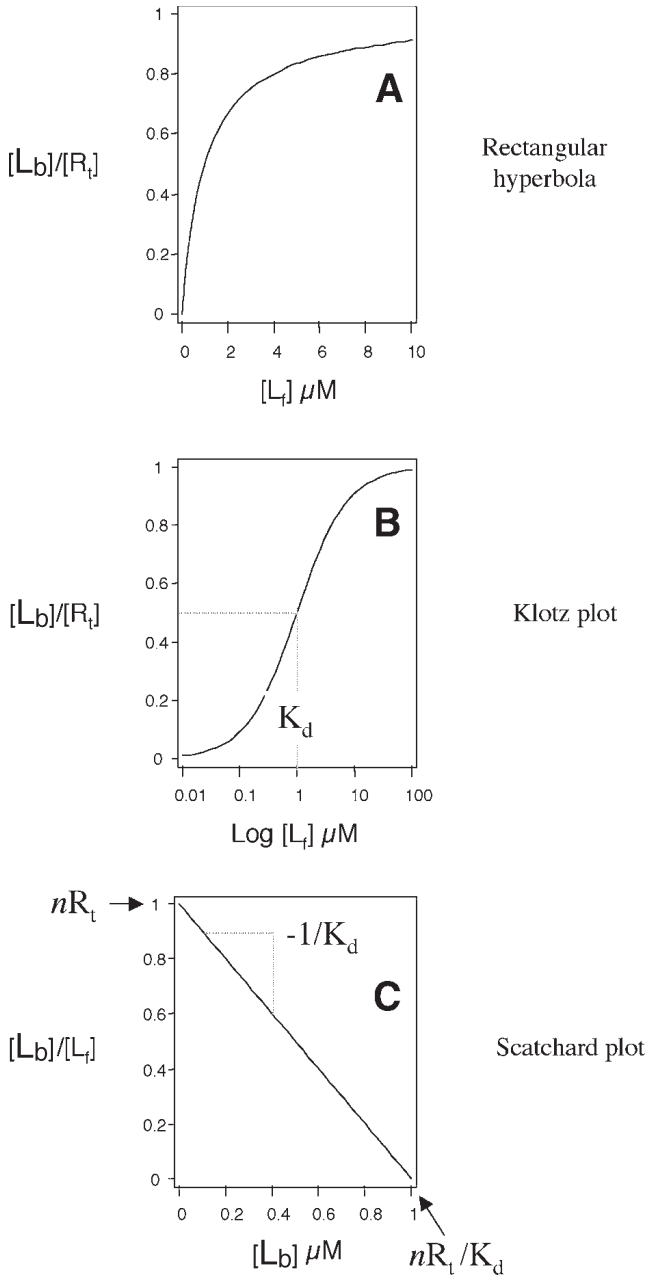


Fig. 1. Plots of simulated data for simple binding. In all cases $n = 1$ and $K_d = 1 \mu M$. **(A)** Direct plot of fractional saturation vs free ligand; **(B)** Klotz plot of the same data, note the log scale; **(C)** Scatchard plot of the same data. The parameters nR_t and nR_t/K_d are estimated from the intercepts.

solubility or sample amounts may prevent us from attaining concentrations necessary to reach the plateau.

An alternative way to plot the data is with a Scatchard plot. For the last 30 yr this has been the traditional method for the analysis of binding data where $[L_f]$ is measured. The Scatchard plot is described by

$$\frac{[L_b]}{[L_f]} = \frac{-[L_b]}{K_d} + \frac{n[R_t]}{K_d} \quad (7)$$

In the simple model, a plot of ligand bound vs ligand bound/ligand free gives a straight line with the x -intercept = $n[R_t]$, a y -intercept of $n[R_t]/K_d$, and a slope of $-1/K_d$ (**Fig. 1C**) (5).

Before the advent of computers, estimates of K_d and n were obtained by any of a number of transformations of the relevant equations to give linear plots. These transformations included the double reciprocal plot and the Scatchard plot. These linearizations are notoriously difficult to fit and generally fraught with problems. *The preferred method of obtaining K_d and n from binding data is direct fitting of the data using a nonlinear least squares fitting algorithm.* Many commercial packages for doing such fits are available today. If we do not explicitly measure the concentration of ligand free, an appropriate solution of the binding equation to obtain the dissociation constant requires that we determine and fit the fractional saturation as a function of the concentration of total L added. The solution of the equation for $[RL]/[R_t]$ as a function of $[L_t]$ is a quadratic equation with the following real solution:

$$\frac{[L_b]}{n[R_t]} = \frac{([L_t] + n[R_t] + K_d) - \sqrt{([L_t] - n[R_t] - K_d)^2 - 4[L_t]n[R_t]}}{2n[R_t]} \quad (8)$$

1.3.2. Nonidentical Binding Sites

Although the most common reason for observing multiple nonidentical binding sites in a protein–protein interaction is likely to be nonspecific binding (see below), it is always possible that there are two independent and noninteracting sites with different affinities. Either case will manifest itself as a deviation from the expected behavior for a simple binding model. The Scatchard plot is a useful diagnostic tool to point out such deviations (**Fig. 2**). A Scatchard plot that is concave upward is indicative of nonspecific binding, negative cooperativity, or multiple classes of binding sites. A concave downward plot suggests either positive cooperativity or instability of the ligand. In any case, proper analysis of this behavior requires other information (for instance, stoichiometry or stability), and the data are best fitted using nonlinear least squares fitting of the data according to an appropriate model.

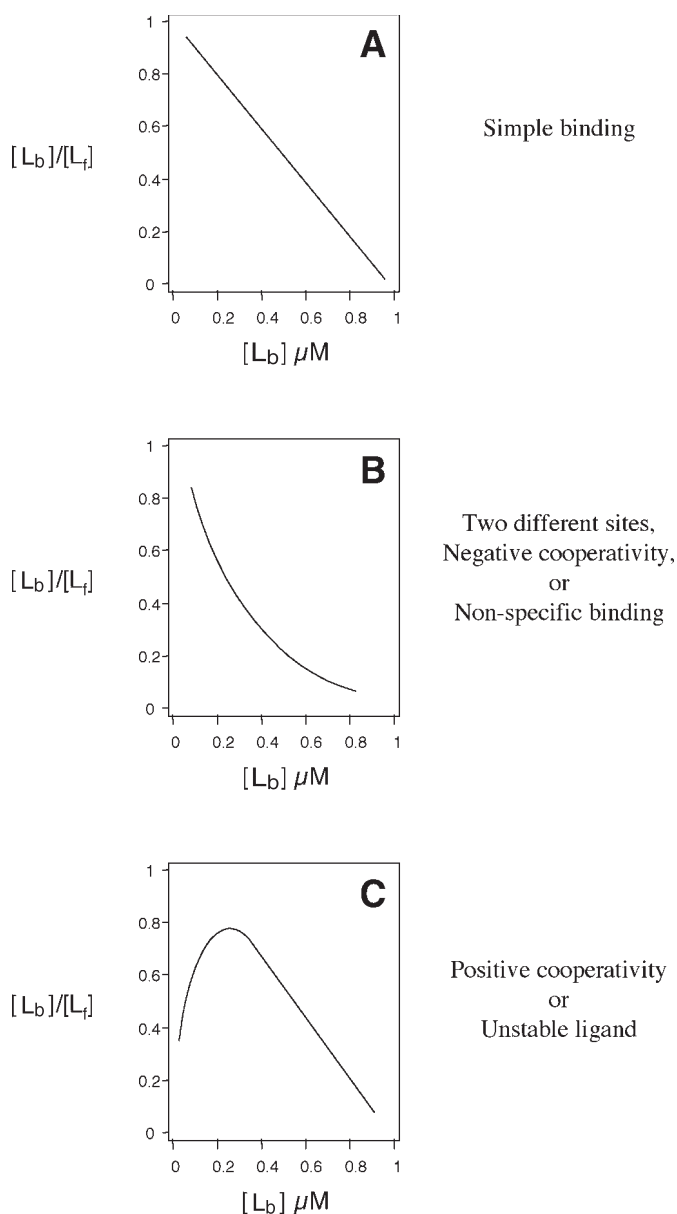


Fig. 2. Effects of complexities on the appearance of the Scatchard plot. (A) Represents the expected behavior in the simple case; (B) a concave upward deviation as shown in this panel could be caused by the presence of two different sites, the presence of negative cooperativity, or a significant nonspecific binding component; (C) Positive cooperativity or ligand instability would lead to the curvature shown in this panel.

Most deviations from simple binding are expected to be due to either multiple sites or nonspecific binding, which, as discussed below, may be difficult to distinguish (8). Either case can be fitted with appropriate modifications of the simple binding expressions. Note that a satisfactory analysis of such complicated binding will require measurement of $[L_f]$.

1.4. Cooperativity

Cooperativity is the term used to describe the situation where occupancy of one site changes the affinity for ligand at another site. There have been many treatments of cooperative binding interactions, including analysis by Scatchard and Hill plots, but these are beyond the scope of this discussion. In general, models explaining cooperativity invoke subunit–subunit interactions in oligomeric protein structures and may well be important in cases where multiple proteins are being assembled into a multimeric complex. The reader is referred to any of several other treatments of such binding if complications of this sort are indicated (9–11). However, it may be simpler to restrict the measurements to conditions where individual subcomplexes are assembled at saturating concentrations before measuring the binding of a subsequent protein.

2. Materials

The only materials relevant to this chapter are a computer and a program to mathematically fit the data. Many commercial and shareware packages capable of nonlinear fitting of equations are available for all platforms; i.e., Prism (GraphPad Software, Inc., San Diego, CA), SigmaPlot (SPSS Science, Chicago, IL), Mathematica (Wolfram Research Inc., Champaign, IL), and DynaFit (BioKin, Ltd., Pullman, WA). There are also published solutions using the popular spreadsheet Microsoft Excel (12,13). The choice is largely a personal preference.

3. Methods

Several chapters in this book describe techniques for determining fractional saturation and/or binding parameters. These basically fall into two categories: direct methods that measure the actual concentration of bound or free ligand and indirect methods that infer the concentrations from some measured signal. The choice of which technique to use may be limited by the strength of the interactions and the inherent sensitivity of the technique. For instance, NMR may be a poor choice to monitor binding constants tighter than micromolar since one commonly needs millimolar concentrations of protein to see a signal. Thus, $[R_f]$ may be $\gg K_d$, and we would be restricted to measuring only the stoichiometry under these conditions (see below). Similarly, with an interaction of millimolar affinity it may be difficult to determine the stoichiometry,

because it may not be possible to attain a concentration of $[R_t] \gg K_d$. See below for a discussion of the relationships between K_d and $[R_t]$.

3.1. Direct Measurement of Free Ligand

Direct methods require that we accurately determine the concentrations of free and bound ligand. Examples of techniques that yield such information include gel filtration, ultracentrifugation, ultrafiltration, or equilibrium dialysis. For binding with slow dissociation rates pull-downs, band shift, or electrophoresis techniques *may* be appropriate. If the process of separating the bound and free ligand is fast compared to the rate of dissociation of the complex, such methods can yield directly the concentrations of bound and free ligand. If dissociation and separation of bound and free reactants occur on similar time scales, such methods are not appropriate for quantitation as the equilibrium will be disturbed by the separation of the reactants. For the same reasons techniques such as cross-linking may overestimate the concentration of RL as the removal of RL will disturb the equilibrium.

3.2. Indirect Measurements of Bound Ligand

More commonly, an indirect measure of saturation is used to monitor binding. These include optical methods such as fluorescence, absorbance, and resonance techniques. These methods all assume that the output signal is directly proportional to the concentration of RL present. For instance, if a fluorescence change is being monitored, it is assumed that there are only two states, the bound and the free, and that each has a characteristic value. If S_0 is the signal in the absence of binding, S_L the signal in the presence of total ligand concentration L , and S_∞ is the value at saturation, then

$$\text{fraction saturation} = \frac{S_L - S_0}{S_\infty - S_0} \quad (9)$$

The concentration of free ligand can be calculated by assuming a stoichiometry n and using the expression $[L_f] = [L_t] - n[R_t]$. Note that if n is incorrect, then the calculated $[L_f]$ will be incorrect also and this will be apparent in the deviation of the data from the theoretical rectangular hyperbola. This is one reason why the determination of n is an important exercise in most binding studies. Alternatively, and preferably, data are fitted using nonlinear least squares methods and n is determined directly from this analysis.

3.3. Competition Methods

Direct methods measure either bound ligand $[RL]$ or free ligand $[L_f]$ as a function of $[L_t]$, and indirect methods usually involve measuring fractional saturation $[RL]/n[R_t]$ as a function of $[L_t]$. However, one of the most useful

variations of the binding experiment is the use of competitive binding assays where a single labeled indicator ligand can be bound and subsequently displaced by any of a variety of competitive inhibitors (**14–19**). Such experiments are particularly useful if the affinity of a series of inhibitors is to be determined. Methods such as fluorescence depolarization or fluorescence resonance energy transfer are particularly well suited for such measurements. A small amount of the labeled ligand is first bound to the receptor and subsequently displaced by titrating with unlabeled inhibitor. The K_i of the unlabeled inhibitor is then calculated. The labeled ligand does not have to be physiological or bound with a physiological affinity because we are always comparing the K_i of the unlabeled inhibitor. Thus, any adverse effects of labeling the indicator ligand will be unimportant.

The IC_{50} is the concentration of inhibitor necessary to displace half the labeled ligand. If $[R_t] \ll K_d$, IC_{50} is related to K_i , the affinity of the unlabeled ligand by

$$K_i = \frac{IC_{50}}{1 + L_t/K_d} \quad (10)$$

where $[L_t]$ is the concentration of labeled ligand and K_d is its dissociation constant. If only relative affinities are to be measured, then comparing IC_{50} directly is sufficient. If absolute affinities are desired, then we must also determine the concentration and affinity of the labeled ligand in the assay.

If $[R_t]$ is similar to or greater than K_d and/or K_i , it follows that the concentrations of free ligand and inhibitor are not equal to their respective total concentrations. For this reason, it is simplest to work at conditions where $[R_t] \sim 0.1 \times K_d$ so that less than 10% of the labeled ligand is bound to the receptor at the start of the experiment.

If higher concentrations of receptor are necessary or if inhibitor binds much tighter than ligand, then one has to fit with a more complex equation (**6,15,17–19**). The following treatment was first published by Wang in 1995 (**19**) and is suitable for fitting the data from competitive displacement experiments where absorbance, fluorescence, or fluorescent anisotropy are measured using commercially available fitting programs. Consider, for example, the binding of a fluorescent probe *A* to a nonfluorescent protein *P* in the presence or absence of a competitive inhibitor *B* that prevents binding of *A*.

Given: $K_a = [A_f][P_f]/[PA]$

$$[A_f] + [PA] = [A_t]$$

$$[P_f] + [PA] + [PB] = [P_t]$$

$$K_b = [B_f][P_f]/[PB]$$

$$[B_f] + [PB] = [B_t]$$

then eq. (11) describes the fractional saturation:

$$\frac{S - S_0}{S_\infty - S_0} = \frac{2\sqrt{(a^2 - 3b)} \cos(\theta/3) - a}{3K_a + [2\sqrt{(a^2 - 3b)} \cos(\theta/3) - a]} \quad (11)$$

where

$$\theta = \arccos \frac{-2a^3 + 9ab - 27c}{2\sqrt{(a^2 - 3b)}^3}$$

$$a = K_a + K_b + [A_t] + [B_t] - [P_t]$$

$$b = K_b ([A_t] - [P_t]) + K_a ([B_t] - [P_t]) + K_a K_b$$

$$c = -K_a K_b [P_t]$$

The experiment requires the measurement of the fractional saturation at various concentrations of A_t , B_t , and P_t . Only a small range of measurements are useful: the ones where fractional saturation is >0.05 and <0.95 . Fractional saturation of P with the probe A is determined by indirect measurements where it is the fluorescence or the anisotropy of AP that gives rise to the signal. The usual experiment is to measure the full binding curve, i.e., $(S - S_0)/(S_\infty - S_0)$ as a function of P_t . This experiment should then be repeated at three or more concentrations of B_t to calculate K_b . Although this may seem like its only giving you three data points, if the curve is fitted, the actual number of useful data points is equal to the total measurements made where fractional saturation is in a useful range.

3.4. Parameters of Reversible Binding

3.4.1. Stoichiometry

Quantitation of binding often requires accurate estimates of the binding stoichiometry n . Many methods are appropriate for this purpose including cross-linking, pull-downs, and electrophoretic methods (when off rates are slow). If association and dissociation rates are fast, these techniques will perturb the equilibrium and give erroneous results. In these cases stoichiometry must be determined from more conventional titrations measuring the equilibrium amounts of RL . To determine stoichiometry an excess of ligand is present and one of the components must be present at concentration well above the K_d in order to ensure saturation. Often this is the first experiment that is done as it helps greatly in fitting the data to more complete titrations.

3.4.2. Kinetics

The analysis of binding requires that we conduct the measurements after binding has reached equilibrium or that we measure individually the rate constants involved. The binding constant can then be calculated from the relationship $K_d = k_2/k_1$. From a practical standpoint, ensuring that the reaction has reached equilibrium often involves measuring a time course for binding at low ligand concentrations and making all measurements after sufficient time to allow attainment of equilibrium. Several examples of each type of analysis are given in subsequent chapters.

In any case, it is instructive to consider the magnitudes of association and dissociation rates. The association rate constants expected for protein–protein interactions are limited by diffusion. If we assume reasonable numbers for the diffusion rate of an average protein, the diffusion limit in aqueous solution is around 10^8 – $10^9 \text{ M}^{-1}\text{s}^{-1}$. There are also additional steric constraints, as only a fraction of the collisions occurring at this rate are oriented properly, and it is commonly assumed that the rate limiting association rate (k_1) for two proteins binding to each other is around $10^8 \text{ M}^{-1}\text{s}^{-1}$.

It can be shown that the rate of approach to equilibrium is determined by the sum of the association rate and the dissociation rate constants. Furthermore, the concentrations of reagents must be at or near the binding constant for accurate determination of both stoichiometry and affinity in the same experiment (*see below*). If the dissociation constant (K_d) for such an interaction is moderate (10^{-6} M), then the dissociation rate for such a complex will be $k_2 = k_1 \times K_d = 10^2 \text{ s}^{-1}$. Thus binding will be complete in seconds and the half-life of the bound state will be tens of milliseconds. If, however, the binding constant is very tight, as may occur in antibody–antigen interactions, the overall equilibrium may take some time. Consider a binding interaction with a free energy of -16 Kcal/mole , an affinity exhibited by many antibodies and other protein–protein interactions (20). This represents a dissociation constant of 10^{-13} M . Here, binding may take as long as hours and the half-life of the bound state could be as long as 20 h. The latter fact is the reason that tight binding can be detected using techniques like immunoprecipitation and pull-down experiments, but tight binding complicates the determination of accurate binding constants.

3.5. Concentrations of Components to Use

3.5.1. Ligand Concentration

Equation (6) is the equation for the familiar rectangular hyperbola with a horizontal asymptote corresponding to 100% saturation and half-maximal saturation occurring at $L_f = K_d$. This equation points out that *the concentrations of*

free ligand present must be similar to the dissociation constant in order to vary the fractional saturation of receptor, i.e., to measure the strength of binding. The most common form of the experiment, then, is to titrate a fixed amount of receptor with variable amounts of ligand and to fit the experimental data to the appropriate binding equation to determine the stoichiometry n and the binding constant K_d .

3.5.2. Receptor Concentration

If we consider the concentration of the fixed protein in this binding equation, i.e., $[R_t]$, we can define three limiting conditions; $[R_t] \ll K_d$, $[R_t] \gg K_d$, and $[R_t] \sim K_d$. **Figure 3** illustrates the interrelationships between K_d and $[R_t]$ in such experiments.

3.5.2.1. $[R_t] \ll K_d$

Under these conditions saturation is achieved by varying $[L]$ at concentrations from 0.1 to 10 times K_d . Because $[L_t]$ is always much greater than $[RL]$ under these conditions, then $[L_f] \sim [L_t]$. Thus, eq. (6) can be simplified to give

$$\frac{[L_b]}{n[R_t]} = \frac{[L_t]}{K_d + [L_t]} \quad (12)$$

If we only measure the fractional saturation (i.e., the ratio $[L_b]/n[R_t]$) as a function of L_t , then we cannot calculate $[L_b]$ because $[L_b] = [L_t] - [L_f]$ and we have not measured $[L_b]$. Note that even if we use direct methods and measure free ligand concentration, the calculation of bound ligand is subject to large errors because the bound is the difference between total and free and, under these conditions, they are about equal (6,7). *Thus, under these conditions, we can accurately determine K_d but not n .* Determination of accurate values for n requires that the concentration of R_t be similar to or larger than K_d .

3.5.2.2. $[R_t] \gg K_d$

If the concentration of R_t is much greater than K_d , then eq. (6) can be rearranged to give

$$\frac{[L_b]}{[L_f]} = \frac{n[R_t]}{K_d + [L_f]} \quad (13)$$

In the first part of the titration curve, when $[L_f]$ is less than K_d (and much less than $n[R_t]$ in this example), the ratio of bound/free ligand is determined solely by the ratio of $n[R_t]/[K_d]$.

If we only measure $[L_t]$, the limiting slope for a plot of saturation vs $[L_t]$ is $n[R_t]/[K_d]$. For example, if $n[R_t]/[K_d] = 100$, then only about 1% of the added

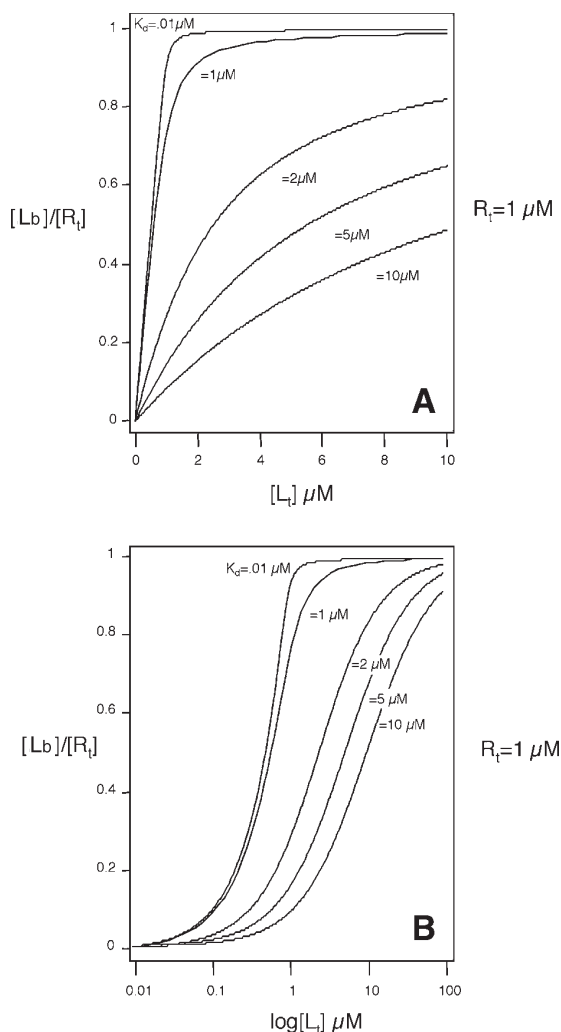


Fig. 3. Binding isotherms for a simple binding equilibrium where $n = 1$ and the total concentration of receptor is $1 \mu\text{M}$. (A) Direct plot of fractional saturation vs total ligand added and (B) the same data plotted on a log scale. Note that as the K_d approaches $[R_t]$ there is a significant deviation from the rectangular hyperbolic behavior.

ligand is free at low ligand concentrations. In order to saturate binding, $[L_t]$ must exceed $100 \times K_d$. When $[R_t] \gg K_d$, the saturation curve is really an end-point determination consisting of two lines (first a slope of approx n and then 0 intersecting at $[L_t] = n[R_t]$) with little curvature (**Fig. 3**). Under these conditions we can accurately determine n , but not K_d .

If direct methods to measure free ligand are used, we can, in theory, calculate K_d , but in practical terms the curve will only deviate from its biphasic nature near $L_t = R_t$ and generally there will not be enough data in this region to obtain accurate estimates of K_d .

3.5.2.3. $[R_t] \sim K_d$

The most useful conditions for determining both K_d and n are when $[R_t] \sim K_d$. The binding curve still resembles a rectangular hyperbola but with small deviations due to the fact that $[L_t] = [L_f] + n[RL]$. Because $[L_f]$ is similar in magnitude to $[L_b]$, each can be measured (or calculated) with good accuracy. *Under these conditions we can determine both K_d and n with a good degree of accuracy from the same experiment.*

4. Notes

4.1. Nonspecific Binding: Specificity vs Affinity

Almost any real-life binding experiment will show some low-affinity binding that is often attributed to “nonspecific binding.” If indirect methods are used to monitor binding, one may or may not see this binding step and one must evaluate if the technique being used will reveal nonspecific binding (i.e., does the detection of binding require occupancy of a specific site such as in fluorescence resonance energy transfer techniques). Nonspecific binding usually presents as an additional slope added to the familiar rectangular hyperbola apparent at high ligand concentrations and the temptation is to simply subtract the linear phase from the observed binding to obtain the specific binding profile. The ambiguity as to whether this binding is “specific” (but just low affinity) or whether this is “nonspecific” has, and will, bedevil many studies (6,16,21). Numerous hydrophobic and ionic interactions can lead to nonspecific binding, but these may be saturable and show a defined n value when two large proteins are involved. Because the binding may well be saturable, the linear subtraction of nonspecific binding may not be appropriate. If we restrict ourselves to consider only two classes of sites, one tight site binding n_1 molecules with affinity K_{d1} and a second weaker site (either due to another specific site or nonspecific binding) binding n_2 molecules with affinity K_{d2} , then we can modify eq. (5) to give

$$[L_b] = \frac{n_1 [R_t] [L_f]}{K_{d1} + [L_f]} + \frac{n_2 [R_t] [L_f]}{K_{d2} + [L_f]} \quad (14)$$

Direct fitting of the data to this expression will allow assessment of both classes of sites. If $K_{d2} \gg [L_f]$, then the second term is approximately linear with $[L_f]$ and this is similar to the usual case of nonspecific binding. But if $K_{d2} \sim [L_f]$,

then the second term will not be linear. Thus, it is preferable to simply fit the binding as though there are two different but independent binding sites. After the data are analyzed with no assumptions, one can question if this interaction occurs at a defined site and in a physiological range of concentrations and is therefore relevant.

4.2. Curve Fitting and Adequacy of the Models

Deviations from the simple binding expressions indicate complexity such as multiple sites or cooperativity. However, the simplest model that explains the data is to be preferred. If the data fit a model with two independent binding sites no better than that with one, the one-site model should be chosen unless there is independent evidence to suggest two sites. Methods of evaluating the goodness of fitting are beyond the scope of this chapter, but are often provided with available fitting programs and should be evaluated before proposing a more complicated expression (7).

4.3. Procedures and Problems

To summarize, the determination of K_d and n for a protein–protein interaction requires that we select a technique appropriate for the binding affinity to be measured. The best concentration of receptor is near the K_d and the concentration of ligand should be varied from two orders of magnitude below to two orders of magnitude above the K_d . The concentration of bound ligand should be determined as a function of the free ligand and the data should be fit to the simplest appropriate model. Generally, n can be determined with a precision of $\pm 20\%$ and K_d within a factor of 2.

Several experimental limitations and errors can limit the accuracy and correctness of the observed fits. Common problems (<http://www.panvera.com/tech/fpguide/FP7.pdf>) are

1. Incorrect correction for nonspecific binding or additional loose binding sites. The suggested solution is to fit to eq. (14).
2. Pooling data from experiments with different receptor concentrations. This will be a problem if the receptor concentrations are near K_d . To avoid this, collect enough data from each titration to do an independent fit and compare the fitted parameters from independent determinations.
3. Presence of a nonbinding contaminant in the receptor or labeled ligand. This may be relevant when labeling the ligand damages the protein, when recombinant proteins are used, and when there is undetected heterogeneity due to misfolded protein.
4. Use of a labeling method for the ligand that alters the binding behavior of that ligand. Use of truncated constructs or incorporation of epitope tags or fluorescent labels may be particularly troublesome. Such problems may be revealed if one compares the apparent affinity from direct experiments using titration with

labeled ligand to experiments where unlabeled ligand is used to displace labeled ligand.

5. Inadequate number of data points or range of ligand concentrations. This is avoided by collecting enough data points, especially at high ligand concentrations.

References

1. Auerbach, D., Thaminy, S., Hottiger, M. O., and Stagljar, I. (2002) The post-genomic era of interactive proteomics: facts and perspectives. *Proteomics* **2**, 611–623.
2. Pawson, T., Raina, M., and Nash, P. (2002) Interaction domains: from simple binding events to complex cellular behavior. *FEBS Lett.* **513**, 2–10.
3. Johnson, M. L. and Straume, M. (2000) Deriving complex ligand-binding formulas. *Meth. Enzymol.* **323**, 155–167.
4. Klotz, I. M. (1985) Ligand–receptor interactions: facts and fantasies. *Q. Rev. Biophys.* **18**, 227–259.
5. Munson, P. J. and Rodbard, D. (1983) Number of receptor sites from Scatchard and Klotz graphs: a constructive critique. *Science* **220**, 979–981.
6. Swillens, S. (1995) Interpretation of binding curves obtained with high receptor concentrations: practical aid for computer analysis. *Mol. Pharmacol.* **47**, 1197–1203.
7. Motulsky, H. J. and Ransnas, L. A. (1987) Fitting curves to data using nonlinear regression: a practical and nonmathematical review. *FASEB J.* **1**, 365–374.
8. Mendel, C. M. and Mendel, D. B. (1985) “Non-specific” binding. The problem, and a solution. *Biochem. J.* **228**, 269–272.
9. Tuk, B. and van Oostenbruggen, M. F. (1996) Solving inconsistencies in the analysis of receptor-ligand interactions. *Trends Pharmacol. Sci.* **17**, 403–409.
10. Koshland, D. E., Jr. (1996) The structural basis of negative cooperativity: receptors and enzymes. *Curr. Opin. Struct. Biol.* **6**, 757–761.
11. Forsen, S. and Linse, S. (1995) Cooperativity: over the Hill. *Trends Biochem. Sci.* **20**, 495–497.
12. Hedlund, P. B. and von Euler, G. (1999) EasyBound—a user-friendly approach to nonlinear regression analysis of binding data. *Comput. Methods Programs Biomed.* **58**, 245–249.
13. Brown, A. M. (2001) A step-by-step guide to non-linear regression analysis of experimental data using a Microsoft Excel spreadsheet. *Comput. Methods Programs Biomed.* **65**, 191–200.
14. Jezewska, M. J. and Bujalowski, W. (1996) A general method of analysis of ligand binding to competing macromolecules using the spectroscopic signal originating from a reference macromolecule. Application to *Escherichia coli* replicative helicase DnaB protein nucleic acid interactions. *Biochemistry* **35**, 2117–2128.
15. Schwarz, G. (2000) A universal thermodynamic approach to analyze biomolecular binding experiments. *Biophys. Chem.* **86**, 119–129.

16. van Zoelen, E. J. (1992) Analysis of receptor binding displacement curves by a nonhomologous ligand, on the basis of an equivalent competition principle. *Anal. Biochem.* **200**, 393–399.
17. van Zoelen, E. J., Kramer, R. H., van Moerkerk, H. T., and Veerkamp, J. H. (1998) The use of nonhomologous scatchard analysis in the evaluation of ligand-protein interactions. *Trends Pharmacol. Sci.* **19**, 487–490.
18. van Zoelen, E. J., Kramer, R. H., van Reen, M. M., Veerkamp, J. H., and Ross, H. A. (1993) An exact general analysis of ligand binding displacement and saturation curves. *Biochemistry* **32**, 6275–6280.
19. Wang, Z. X. (1995) An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Lett.* **360**, 111–1114.
20. Brooijmans, N., Sharp, K. A., and Kuntz, I. D. (2002) Stability of macromolecular complexes. *Proteins* **48**, 645–653.
21. Rovati, G. E., Rodbard, D., and Munson, P. J. (1988) DESIGN: computerized optimization of experimental design for estimating K_d and B_{max} in ligand binding experiments. I. Homologous and heterologous binding to one or two classes of sites. *Anal. Biochem.* **174**, 636–649.



<http://www.springer.com/978-1-58829-120-2>

Protein-Protein Interactions

Methods and Applications

Fu, H. (Ed.)

2004, XVI, 532 p., Hardcover

ISBN: 978-1-58829-120-2

A product of Humana Press