Biomolecular Ligand-Receptor Binding Studies: Theory, Practice, and Analysis

Charles R. Sanders,
Dept. of Biochemistry, Vanderbilt University

Table of Contents

One of the ways biological chemistry distinguishes itself from traditional chemistry is by the degree to which biochemistry is based on NON-COVALENT and REVERSIBLE binding. This is vital to life. For example, enzymes are the machines that catalyze the chemistry of physiological processes. Enzymes reversibly and non-covalently bind their substrates as part

of the catalytic cycle. The same principle extends to the regulation of biochemical processes. If it is said that a certain ligand directly "regulates" an ion channel, it is generally implied that the molecule *reversibly binds* to the ion channel, causing it to shift towards open or closed states. The entire immune system is based upon the body being able to produce "custom non-covalent binders" for just about any molecule it wants to.

Binding can be also be exploited by humankind The exploitation of natural binding processes lies at the heart of a number of basic techniques in molecular biology and biotechnology. For example, various immunoassays, such as Western blots, are based on antibody binding to signature molecules associated with various biological cells and processes. Endonucleases are exquisite in their ability to distinguish particular nucleotide sequences from all other sequences, leading to specific binding followed by cleavage, and are thereby standard tools of molecular biology.

Drugs are typically molecules that are bound by biomolecules (usually proteins or nucleic acids) with a high degree of affinity and specificity, leading to a medically relevant physiological response. Most drugs interfere with or modulate a target protein through non-covalent binding.

In common lab practice, "affinity chromatography", such as metal ion affinity chromatography, is based on attaching protein or ligand molecules to a gel matrix and then passing a mixture through the column: only the molecules that have a strong binding affinity for the matrix-bound molecule will stick to the column- everything else passes right on through. The now pure molecule on the column is then displaced by passing a solution through the column that competes with the matrix for the specially-bound molecule being purified.

It may be helpful to review some vocabulary with respect to biomolecular association/dissociation (binding):

Ligand and Receptor: For any two non-identical molecules that associate it is possible refer to one as the "Ligand" and to one as the "Receptor". Typically, the receptor will be a protein and the ligand will be a smaller molecule, but not always. "Receptor" for the purposes of this discussion is being used generically does not imply a biological function as a receptor (for example, enzymes and transport proteins will be called "receptors" in this discourse). In fact, one person's "ligand" may be another's "receptor". Consider protein-DNA association. Which is the ligand and which is the receptor? It really doesn't matter, the terms are basically interchangeable. Usually, the receptor will be the larger of the two molecules. The reason for adopting this terminology lies in the fact that in most binding studies, the "receptor" concentration is usually held constant (or nearly so) and the ligand concentration is varied.

Molecular Recognition: Humans recognize objects and people using the 5 senses. In fact, humans have extraordinary abilities to discern. Molecules are somewhat less sophisticated. Molecules "recognize" one another when they come close enough to "feel" the presence of each other- either by physically bumping into each other or through the interactions of the fields (electrostatic) between each other. Usually we use "molecular recognition" in a positive sense, meaning that molecules that are energetically compatible will associate.

Specificity: The "specificity" of a ligand for a receptor (or vice versa) is a description of how favorable the binding of the ligand for the receptor is compared with its possible binding to

other types of receptors that may also be present. For a receptor, “specificity” describes how much the receptor favors a particular ligand relative to the other ligands that may also be present. In real biological systems the specificity of either ligands or receptors is rarely 100%- this is one of the reasons why drugs tend to have side effects. For example, it is well known that most proteins that bind a given nucleotide (like ATP) are not completely specific for ATP, but can bind a variety of ATP analogs like thio-ATP, AMPPCP, or even GTP.

Affinity: "Affinity" simply refers to how strong the binding is (as judged by $K_{association}$ or $K_{dissociation}$ and ΔG^o). "High affinity" refers to very strong binding (large negative ΔG^o and a very small K_d). The association or dissociation constant is often referred to as the "affinity” or “binding” constant.

Stoichiometry: "Stoichiometry" refers to how many molecules of ligand can bind to a single receptor.

Cooperativity: "Cooperativity" (sometimes called “synergism”) refers to situations where the binding of one (or more) molecules to the receptor enhances (or weakens) the binding of additional molecules to that same receptor. Cooperative binding effects are also known as "allosteric effects".

Reversible vs. Irreversible Binding: All non-covalent binding processes are reversible, meaning that the ligand can both bind to and dissociate from the receptor. Equilibrium is reached when the time following mixing is long compared to the $t_{1/2}$ binding and dissociation. However, sometimes non-covalent binding is so tight that the ligand does not dissociate for a very long period of time (sometimes days). In such cases, the association is effectively *irreversible* and does not reach equilibrium within the relevant time frame.

Kinetics: "Kinetics" is a rather generic term used to describe both the rates at which processes occurs and the field associated with the study of rates. Binding and dissociation processes will be characterized not only by the equilibrium constants, but also by how fast association/dissociation occur.

The Simplest Case: 1:1 Stoichiometry

$$R + L \rightleftharpoons RL$$

R: "receptor": could be enzyme, transporter, carrier protein, receptor, etc.

L: "ligand": could be substrate, inhibitor, drug, metabolite, hormone, DNA/RNA, another protein, etc.

It is equally valid to write the equilibrium constant in either of two ways:

$$K_{association} = K_a = \frac{[RL]}{[R] \cdot [L]}$$

Where the concentrations of the free R, free L and the complex are the concentrations at equilibrium

$$K_{dissociation} = K_d = \frac{[R] \cdot [L]}{[RL]}$$

$K_a = 1/K_d$ It is also easy to show that $\Delta G^o_{association} = -\Delta G^o_{dissociation}$

A Short Introduction to Binding Kinetics

Before proceeding to a more detailed consideration of binding theory and analysis, it is important to first understand basic kinetics, a short review of which is presented here. Binding and other equilibrium constants are fundamentally related to the rates of interchange between the states involved in the equilibrium process. "Rate" is, of course, a description of how frequently something happens.

A unimolecular rate is how fast *one* molecule does something and will have units of "per second" (= sec^{-1} = Hz) or "per minute" (= min^{-1}). Unimolecular rates are sometimes referred to as zero order rate constants where "zero" means that the rate is independent of any concentration. One example of a zero order rate constant is the radioactive decay of a single isotope (which is determined completely by the type of isotope, not by chemical concentrations or compositions). Another example is the enzyme turnover number: k_{cat}. This rate constant tells the maximum rate that a single enzyme molecule can execute a chemical reaction under conditions where it is saturated with substrate.

A "first order" reaction rate is a rate that describes a process that is dependent upon the concentration of a single species. It will have units of

$$\Delta[\text{concentration}]/\Delta\text{time}$$

(e.g. mM product produced/minute) or Δquantity/Δtime (e.g. micromoles product produced/second). An example of a first order reaction would be the production of an enzyme-product complex from an enzyme-substrate complex:

In this case the rate of EP production will be dependent upon two things:

i. How much ES is present (its concentration)
ii. A factor that describes how often ES will get converted into EP under standard conditions

The "factor" mentioned above is the first order rate constant, "k". The rate will be defined:

$$\text{rate of EP production} = \Delta[EP]/\Delta t = k \cdot [ES]$$

From this, it can be seen that the units of the first order reaction rate will be $time^{-1}$. Thus, if we define the rate as mMolar of EP produced per minute, the units of k will be min^{-1}. In the above example, if the assumption can be made that ES cannot be converted into anything except for EP, it also would have been possible to define the rate of EP production as the Δ[ES]/Δtime.

Another example of a first order process is the dissociation of a 1:1 receptor-ligand complex to form free ligand and free receptor. While two separate species are produced, the rate at which they are produced will be dependent upon a single concentration: that of the complex.

A "second order reaction" is a reaction whose rate is dependent upon the concentrations of 2 species. For example in the reaction

the reaction rate could be defined as either Δ[C]/Δtime or Δ[D]/time, but the rate will be dependent upon both [A] and [B]:

$$\text{rate} = \Delta[C \text{ or } D]/\Delta\text{time} = k \cdot [A] \cdot [B]$$

where k is the second order rate constant. k is a factor that describes how often A and B react under standard conditions. It can be seen that the units for a second order rate constant must be $concentration^{-1} \cdot time^{-1}$. For example: per molar per second ($M^{-1} \cdot sec^{-}$).

An example of a second order process is the binding of a ligand (such as a hormone) to a receptor (such as a GPCR) to form a 1:1 ligand-receptor complex. In this case, the rate is dependent upon both concentrations: L and R can associate only if they bump into each other and the probability that they will bump into each other is determined by their concentrations.

The Variables of Binding Studies

As we shall see, there is diversity in the classes of binding processes- a tremendous range of possible stoichiometries, variable affinities, cooperativity, etc. There is also a remarkable wealth of techniques for examining binding experimentally. We will therefore selectively survey binding theory and techniques. Towards this end, a logical starting point is overview the possible variables relevant to studies of biomolecular association.

Consider the simple case of the formation of a 1:1 complex:

$$R + L \rightleftharpoons RL$$

where:
$$K_d = 1/K_a = \frac{[R] \cdot [L]}{[RL]} \qquad (1)$$

From this equation it can be seen that $K_{dissociation}$ and $K_{association}$ for a given system can be determined any time the concentrations of [R], [L], and [RL] are measured under equilibrium conditions. This is the basis for one entire class of experimental methods to study binding.

The free energy in favor of binding (negative is favorable) is:

$$\Delta G^o = R \cdot T \cdot \ln K_{dissociation} = -R \cdot T \cdot \ln K_{association} \qquad (2)$$

Note that in addition to being useful as an equilibrium constant, K_d gives the free ligand concentration at which the total populations of free and complexed receptors will be equal (half maximal binding). This is an important fact to know.

Relationship Between Thermodynamics and Kinetics of Binding

The strength of binding is related to the "kinetics" of ligand-receptor association-dissociation: how fast the ligand binds and how fast it dissociates.

As described in the previous section the rate of bimolecular processes is dependent upon the concentrations of the species involved and a rate constant. For a generic 1:1 ligand/receptor binding system:

$$L + R \rightleftharpoons LR$$

The forward rate is $k_{on} \cdot [R] \cdot [L]$, while the reverse rate is $k_{off} \cdot [RL]$. k_{on} is a second order rate constant while k_{off} is a first order rate constant. By definition, at equilibrium the rate of the forward process equals that of the reverse process, meaning:

$$k_{on} \cdot [L] \cdot [R] = k_{off} \cdot [LR] \qquad (3)$$

where the concentrations are equilibrium concentrations. This expression can be rearranged:

$$k_{off} \qquad [L] \cdot [R]$$

$$\frac{}{k_{on}} = \frac{}{[LR]} \qquad (4)$$

which is (of course) equal to the dissociation constant, K_d. From this derivation, it can be inferred that another class of experimental binding study methods is based upon making kinetic measurements (because if the on and off rate constants are determined, then K_d is determined).

The rates are related to the time that the ligand spends in the free and bound environments. For example, the "half-life" describing the average amount of time a ligand will spend as part of a complex is:

$$t_{1/2} = 0.693/k_{off} \qquad (5)$$

This actually can tell us something important. The on/off times will influence the choice of binding methods to study binding in a particular system (see below).

What determines k_{on}? It is usually determined primarily by how fast (how often) the receptor and ligand bump into each other. This is determined mostly by the rate of diffusion of the ligand. For typical sized ligands, k_{on} will fall roughly in the range of 10^6-10^8 per molar receptor per second. This is very fast compared to the off rate under normal conditions.

What will the range of off-rates and "complex" dissociation half lives be? It is K_d-dependent because $K_d = k_{off}/k_{on}$. So, if k_{on} falls in the range of 10^6-10^8 $M^{-}sec^{-}$, we can roughly estimate the following:

Kd (M)	**k_{off} range (sec^{-})**	**$t_{1/2}$ range**
10^{-11}	10^{-5} to 10^{-3}	hours to days
10^{-9}	10^{-3} to 0.1	seconds to hours
10^{-7}	0.1 to 10	0.1 to 10 seconds
10^{-6}	1 to 100	msec to 1 second
10^{-3}	10^3 to 10^5	10 μsec to 1 msec

The take-home message of this table is two-fold. First, when binding is very tight, complexes can persist for a long time (hours). This has implications for design of binding experiments (see below). Secondly, when a ligand and a receptor are mixed, it will take at least 5 times $t_{1/2}$ for equilibrium to be approached. In cases where binding is very strong, this means waiting a very long time! If one can't wait long enough, then for all practical purposes the binding is irreversible.

One caveat that should be mentioned is that there are certain classes of ligand and receptors for which k_{on} is lower than the diffusion rate limit. This is known as "slow binding" behavior and is not rare. For a given K_d, a much slower on rate also means a much slower off rate compared to the values given in the table. Slow binding behavior arises from factors that

cause a lowering of the probability of successful association every time the two molecules bump into each other. Lowering the probability of productive contact means a lower association rate. For example, if a protein can exists in two major conformations, one that binds and one that doesn't, every time its ligand bumps into its inactive form, they will fail to form a complex.

From this section, it can be seen that a measurement of K_d requires at least the measurement of the concentrations of the free and bound species at equilibrium or determination of the on and off rate constants.

The Attractiveness of Studying Binding Using Pure Ligand(s) and Receptor

One fact that should be apparent is that binding studies become much more straightforward if one is working with pure ligand and receptor and under well-defined experimental conditions. Whenever a system is studied *in vivo* (or in a biological extract containing numerous types of molecules), the number of experimental variables becomes very high, to the point that it may be difficult to obtain unambiguous results (see following sections).

The Model for 1:1 Binding

In the very simplest type of binding study involving a simple 1:1 association of ligand and receptor to form a complex, a primary goal might be to determine the K_d.

The standard definition of K_d can be algebraically manipulated to yield the following equation:

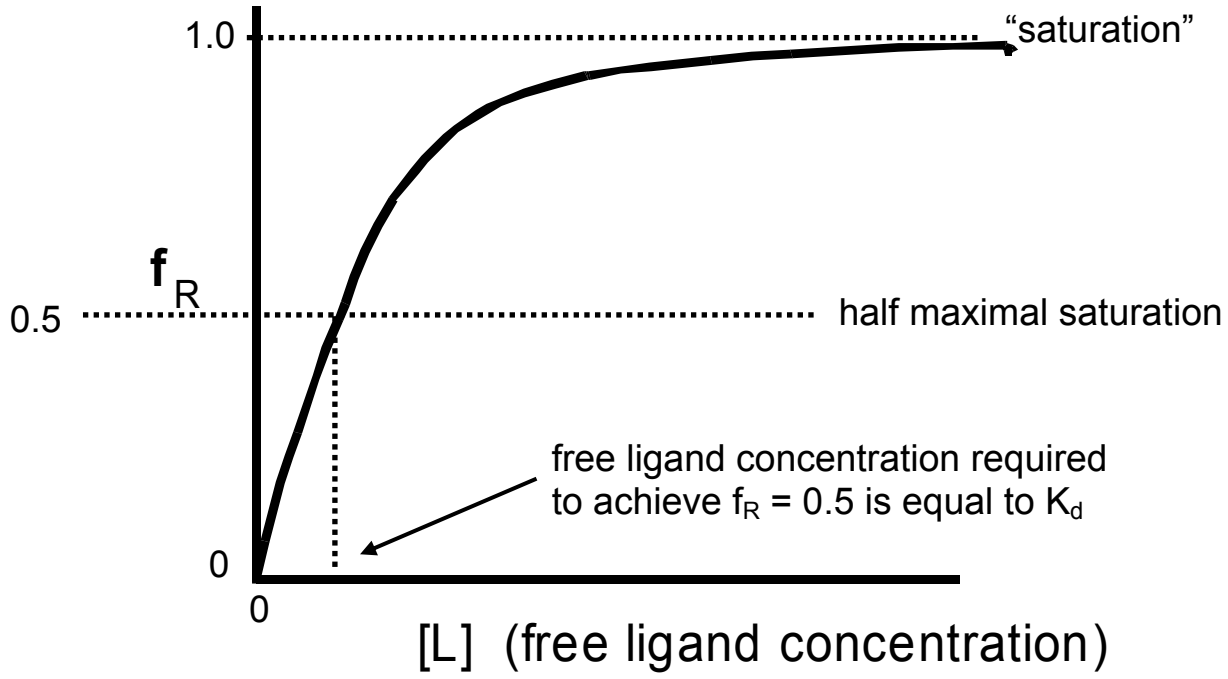
$$[RL]/[R]total = \text{fraction of sites occupied} = f_R = \frac{[L]_{free}}{K_d + [L]_{free}} \qquad (6)$$

$$= \frac{K_a \cdot [L]_{free}}{1 + K_a \cdot [L]_{free}} \qquad (7)$$

where $[L]_{free}$ is the free ligand concentration. From this equation it is observed that if one could measure the fraction of sites occupied as a function of $[L]_{free}$, the data would map out a curve that could be fit to yield a value for K_d.

Equations 6 and 7 predict hyperbolic ($fraction_R$) vs $[L]_{free}$ plots (see below). Such plots are sometimes referred to as "isotherms" (for relatively obscure thermodynamic reasons).

The "binding isotherm" equation for 1:1 binding can be plotted:



Maximum % change in f_R per unit [L] is in the 0-0.5 range

Important implications and considerations for 1:1 binding isotherm:

Remember, *it is free [L] that is being plotted*. However, since total ligand is often much higher than total receptor, this means that the % ligand that forms a complex with R is often going to be small. In this (very common) case $[L]_{free}$ is effectively equal to $[L]_{total}$. This is fortunate, because the total ligand concentration is often easily determined, but not the free ligand concentration. THE ASSUMPTION THAT L_{free} IS EFFECTIVELY EQUAL TO L_{total} CAN OFTEN BE SAFELY MADE, BUT NOT ALWAYS. THIS ASSUMPTION SHOULD ALWAYS BE SCRUTINIZED BEFORE BEING MADE.

To approach 100% saturation requires that [L] be many times higher than K_d. If [L] is a drug there may be negative tradeoffs involved in going to very high drug concentrations in order to achieve maximum efficacy.

When collecting data it is important to choose [L] concentrations both below K_d, at K_d and >> K_d in order to get data from each part of the binding curve… only with thorough representation of all parts of the curve will it be possible to get a reliable fit of the model to the data (that yields a reliable value of K_d) and to verify the applicability of the 1:1 model.

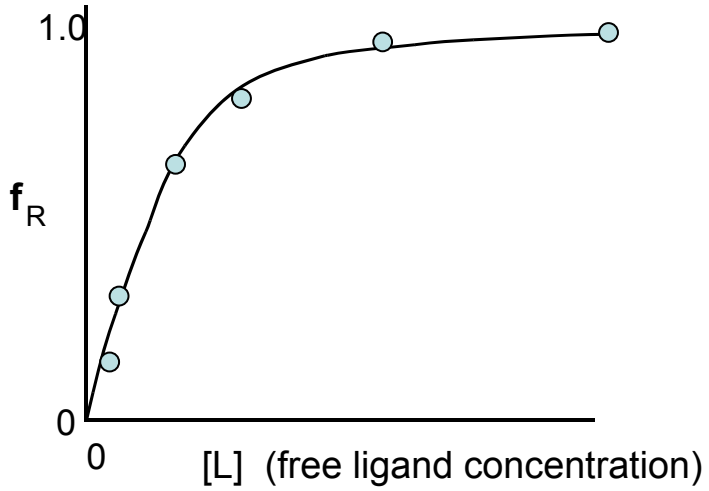
At [L] concentrations up through 1-2 K_d, f_R is very sensitive to ligand concentration.

The isotherm equation (eq. 6) becomes equal to the Michaelis-Menten equation for enzyme kinetics if f_R is replaced with v/V_{max} and K_d with K_m. K_m is sometimes equal to the true K_d, but not always, depending on the enzyme/substrate.

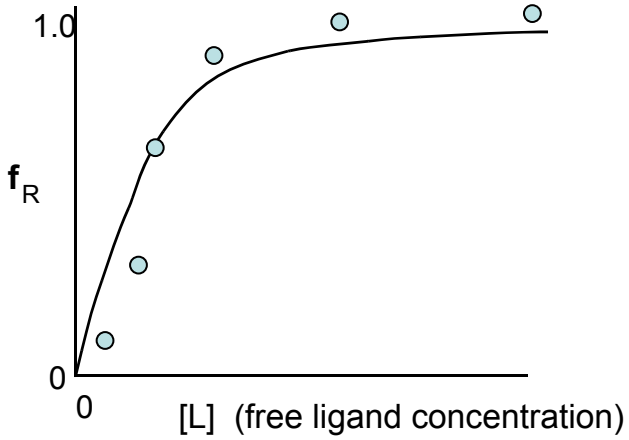
Often, a dose-response may be quantitated in terms of the concentration of ligand required to give 50% maximum effect. Such values are then reported as K_i, LD_{50}, I_{50} or $K_{apparent}$. These values MAY reflect a true K_d, but often do not because a 1:1 binding model may not be valid or because measurements are not being made under equilibrium conditions.

Fitting a Model to Data

If binding is involves 1:1 model, then data should be fit by a hyperbola.



If model is not appropriate, then the model will not be well fit to the data.



Here, the data is clearly reflecting a sigmoidal shape. A 1:1 binding model does not describe to the data adequately. A different model would be more appropriate.

A Little More On 1:1 Binding

The obvious problem of actually using eqs. (6) and (7) is the question of how to experimentally determine the fraction of sites occupied, $[RL]/[R]_t$ at a given composition. This parameter (fraction of sites filled, f_R) is what most experimental binding study methods are designed to yield (see below). For now, let us assume we can find the fraction. Then, how can $[L]_{free}$ be determined? Sometimes it is measured directly (see following sections). However, if $[L]_{total}$ is known in advance (it usually is), $[L]_{free}$ can be determined as follows:

$$[L]_t = [L]_{free} + [RL] \qquad (8)$$

thus: $$[L]_{free} = [L]_t - [RL]$$

and: since fraction of R complexed = $f_R = [RL]/[R]_{total}$ and $[RL] = [R]_t \cdot f_R$, then

$$[L]_{free} = [L]_t - f_R \cdot [R]_{total} \qquad (9)$$

Accordingly, $[L]_{free}$ can be determined if we know the fraction of sites occupied, $[L]_t$, and $[R]_t$.

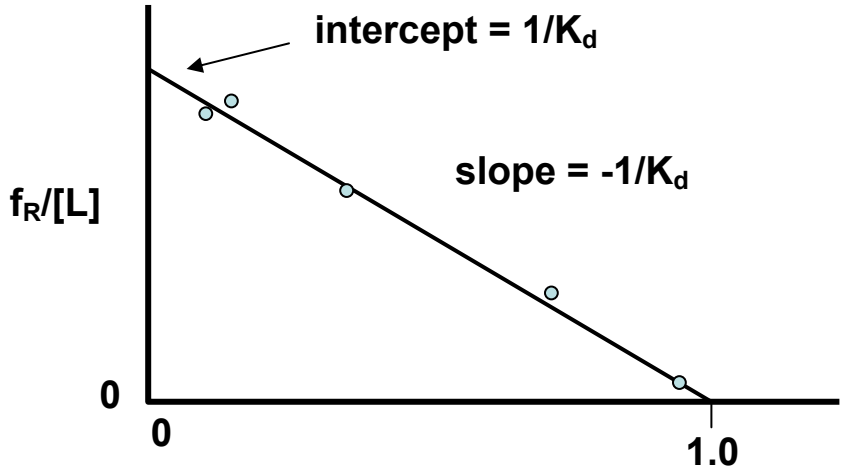
What if $[R]_t$ is not known, as might be the case in studies involving a crude biological extract? If the assumption can be made that $[RL] << [L]_t$, then we can make the approximation that $[L] = [L]_t$. This is often the case in real systems. If not, then we may have to use equation (9) but treat $[R]_t$ as an additional unknown constant that must determined during the analysis.

Based on the above, it can be seen that binding studies can generally be reduced to measuring the fraction bound as a function of the free ligand concentration and possibly the total receptor concentration.

It should be emphasized that there are other ways of rewriting equations 6 and 7. The Scatchard equation is an example of a linear form:

$$\frac{f_R}{[Ligand]} = \frac{1}{K_d} - \frac{f_R}{K_d} \qquad (10)$$

A Scatchard plot for data that satisfies the 1:1 binding model is shown below.



A variation of the Scatchard plot/equation given above is the closely related equation that leads to a plot where $[RL]/[L]_{free}$ is the y variable and [RL] is the x variable.

If we wished, we could rewrite equations 6-7 so that the function is dependent upon the fraction of total ligand that is complexed rather that the fraction of receptor:

$$f_L = \text{fraction of L complexed} = \frac{[RL]}{[L]_t}$$

so: $[RL] = f_L \cdot [L]_t$

which can be substituted for [RL] in equation 6 so that:

$$\frac{f_L \cdot [L]_t}{[R]_t} = \frac{[L]_{free}}{K_d + [L]_{free}}$$

which is easily simplified using algebra to:

$$f_L = \frac{R_t \cdot [L]_{free}}{([L]_t \cdot (K_d + [L]_{free})}$$

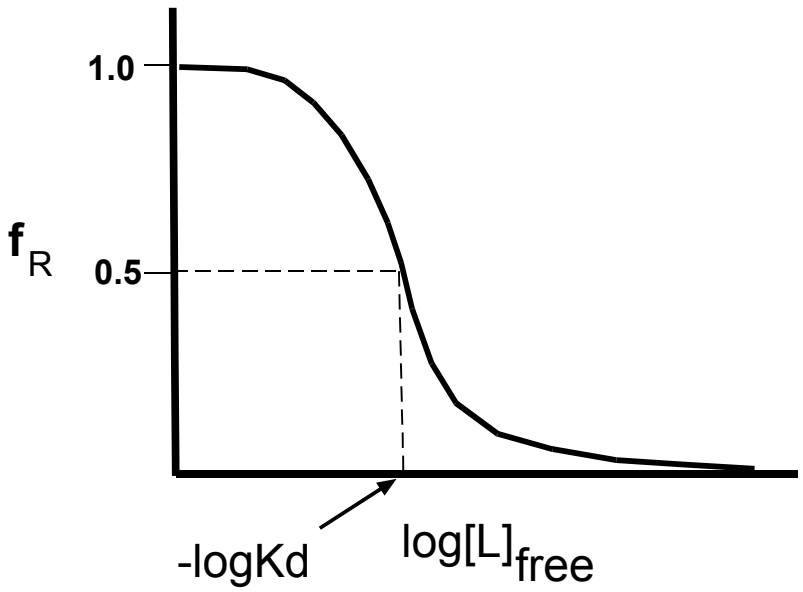
which could be rearranged into a number of linear forms (including Scatchard). From this equation, it can be seen that K_d could also be determined by monitoring the fraction of [L] that is complexed as a function of $[L]_{free}$ and $[R]_t$.

Alternatives to "Direct" and Scatchard Plots

In addition to determining parameters using the direct (eqs. 6-7) and Scatchard-type plots (equation 12) there are a number of other types of plot that are often encountered. For the case a single ligand and a single receptor site here are some examples:

"Bjerrum" Plot:

$-\log[L]_{free} = -\log K_d + \log[(1-f_R)/f_R]$

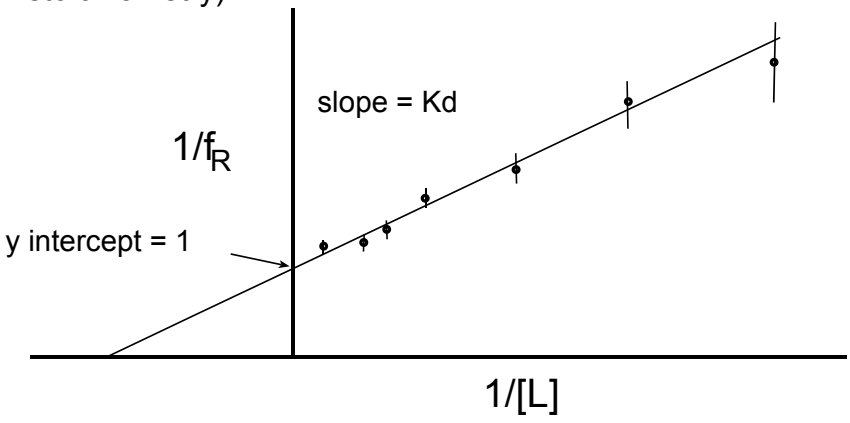


As shown, the predicted curve is sigmoidal, with the point of inflection giving a simple read out of $-\log K_d$. Such plots are often used for acid-base titrations because for such titrations $-\log K_d = pK_a$.

"Double Reciprocal" or "Benesi-Hildebrand" Equation/Plot::

$$\frac{1}{f_R} = 1 + \frac{K_d}{[L]_{free}}$$

That this equation is similar to the Lineweaver-Burke plot, where f_R corresponds to the observed enzyme reaction rate, K_d corresponds to K_m, $[L]_{free}$ corresponds to the substrate concentration, and "V_{max}" corresponds to the "maximum fraction bound" which is, of course, simply 1.0 (for 1:1 stoichiometry).



Measuring Concentrations in Binding Studies

Ligand. This is often straightforward. The ligand is usually titrated into a solution containing the receptor. The concentration of the stock solution is known either because a precise amount of ligand was weighed out and dissolved in a known quantity of water or because of its specific radioactivity, its UV absorbance, or some other method. Thus, based on the

known stock solution and the amount added to a known volume of receptor solution, the total ligand concentration is easily calculated (it's just a dilution problem).

Receptor. In some, but not all cases, there are straightforward ways to determine the concentration of a macromolecule:

Depending on the method for determining the fraction of receptor sites filled by ligand, it is sometimes possible to determine the receptor concentration in the course of the actual binding study. For example, if it is possible to measure the fraction of ligand that forms a complex, it is easy to determine $[L]_{free}$, such that the fraction ligand complexed (f_L) versus $[L]_{free}$ data can be fit to yield both K_d and $[R]_t$.

Occasionally, the "receptor" can be weighed out (assuming it is pure and available in mg quantities).

If R is pure and is a protein a couple of common protein assays (in cases where it is not available in dried form or when there is not enough to directly weigh out) are:

1. Colorimetric assay (such as the Lowry and Bradford assays): a dye is added to the protein solution that changes colors (its light absorbance properties) upon binding to the protein. The amount of color change (as determined by UV spectrophotometry) is directly proportional to the protein concentration.

2. UV absorbance: many proteins absorb light strongly at 280 nanometers, due to Tyr and Trp. For proteins of known sequence, the extinction coefficient (how much light at a give wavelength gets absorbed from a protein solution of a given concentration) can be calculated. Based on this, one can simply measure A_{280} and then calculate the protein concentration using Beer's law. UV absorbance at 259 nm is a standard way to measure the concentration of nucleic acids (RNA/DNA).

If the receptor is not pure (or anywhere close), it can be difficult to directly determine the concentration. A couple of less direct ways sometimes used are:

1. Measure the "activity" of the protein. For a pure enzyme that, when pure, catalyzes a reaction with a specific activity of 1000 units/(min-mg), if a cell extract shows an activity of 1 unit/ml, then it can be estimated that there is 1 microgram of enzyme per ml of extract. Another "activity" that might be used in quantitation would be (in the case of a cellular membrane receptor) a measurable physiological outcome whose magnitude is directly proportional to the amount of the receptor (does/response).

2. Affinity-based assay. A tagged antibody (or some other molecule) that specifically binds only the receptor of interest is mixed with the receptor solution followed by quantitating how much complex forms. For example, if the receptor is found on a cell surface it may be able to take a gold-labeled antibody solution and incubate it with the cells. The cells can then be washed to remove free labeled antibody and electron microscopy can be carried actually visually count the number of gold "dots" appearing on the cell surface.

What if the Receptor/Ligand System is More Complicated than 1:1 Binding?

Many systems (e. g. proteins involved in regulation, signal transduction, etc.) involve receptors that have more than one of binding site for one or more of ligands. Even when all sites are for the same ligand, they may not all have identical affinities and their affinities may vary depending on whether the other sites are filled or not. These complications of greater-than-1:1 stoichiometry (for a single ligand), binding of more than one type of ligand, non-identical sites, and/or cooperative sites make experiments harder to design and correctly interpret. Let's try to sort out the adjustments that have to be made in such situations.

Multiple Sites for a Single Ligand, K_d is Same for All Sites Regardless of the Current Number of L bound.

$$RL_{0 \text{ to } n-1} + L \rightleftharpoons RL_{1 \text{ to } n}$$

where n is the number of equal affinity binding sites on R for L.

In this case:

$$\text{fraction}_R = \frac{[RL] + 2\cdot[RL_2] + 3\cdot[RL_3] + \ldots n\cdot[RL_n]}{[R]_{tot}} \qquad (11)$$

$$= \frac{n \cdot [L]_{free}}{K_d + [L]_{free}} \qquad (12)$$

which leads to a linear Scatchard equation:

$$\frac{\text{fraction}}{[L]_{free}} = n/K_d - \text{fraction}/K_d \qquad (13)$$

Note that as defined in eq. 13, the maximum "fraction" is not longer equal to 1.0, but is equal to n. The shape of the direct fraction vs. [L] curve will be exactly the same as for a 1:1 complex (a hyperbola).

The Case Where There are Multiple Independent Sites for the Same Ligand, but they Have Differing Affinities.

In this case the following equation holds:

$$\text{fraction}_R = \frac{[RL] + 2\cdot[RL_2] + 3\cdot[RL_3] + \ldots n\cdot[RL_n]}{[R]_{tot}}$$

$$= \sum_{n} \frac{[L]_{free}}{[L]_{free} + K_{d,n}} \qquad (14)$$

In the more general case where there are n types of sites (each with its own K_d) and m of each type, the more general equation is:

$$= \sum_{n} \frac{m \cdot n \cdot [L]_{free}}{[L]_{free} + K_{d,n}} \qquad (15)$$

A fraction vs. [L] binding curve will be the sum of hyperbola. By the naked eye, it may be hard to distinguish the observed curve from a true hyperbola.

This equation can be converted into a Scatchard-type equation:

$$\frac{\text{fraction}}{[L]_{free}} = \sum_{n} \frac{m/K_{d,n}}{1 + [L]_{free}/K_{d,n}} \qquad (16)$$

$$= \sum_{n} (m/K_{d,n} - \text{fraction}/K_{d,n})$$

Now, however, the plot will no longer be linear but will exhibit a curve.

The Case where 1 Ligand Binds to Multiple Receptors

In the above case, it was implied that while there were multiple sites present, they were all on the same receptor (and will have the same concentrations, assuming equivalent stoichiometries).

What about the more general case where there is more than one receptor? In such a case, the fraction of receptor sites bound will not usually be measured (since there will be multiple receptors). Instead one would usually measure the total fraction of ligand bound (to all receptors). In this case, the relevant equation is:

$$\text{fraction of L bound} = \sum_{n} \frac{m_i \cdot [L]_{free}}{K_{d,i} + [L]_{free}} \qquad (16)$$

where n is the number of receptors and m represents the number of identical sites on the "ith" (where i = 1, 2 ... n) receptor.

The Case Where there are m Equivalent Binding Sites That are Homocooperative

$$R + L \rightleftharpoons RL + L \rightleftharpoons RL_2 + L \rightleftharpoons RL_3 \text{ ..... } RL_n$$

$K_{d,1}$ $K_{d,2}$ $K_{d,3}$

Sometimes (quite frequently in biology) a receptor will have multiple binding sites for a ligand that are initially equivalent. However, when a ligand binds to one site it causes a change in the receptor (typically a conformational change) that causes the affinity of the remaining sites for the ligand to either increase ("positive" cooperativity) or decrease ("negative cooperativity"). Binding of a second molecule of the same ligand may induce yet another perturbation and so forth. The exact model for such a situation is:

$$f_R = \frac{K_1 \cdot L_f + 2 \cdot K_1 \cdot K_2 \cdot L_2 + \ldots n \cdot K_1 \cdot K_2 \ldots K_n \cdot L_n}{1 + K_1 \cdot L_f + K_1 \cdot K_2 \cdot L_2 + \ldots + K_1 \cdot K_2 \ldots K_n \cdot L_n} \qquad (18)$$

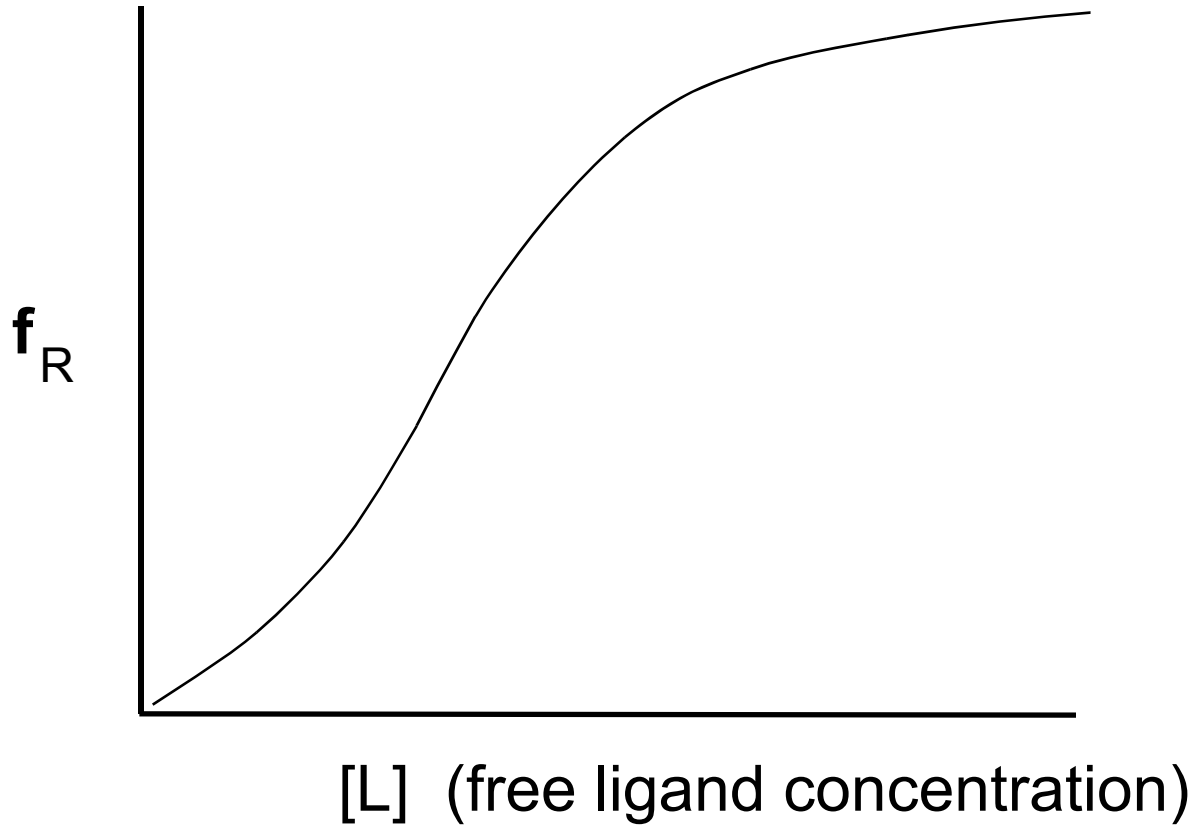
where the K above are *association* constants (remember, these = $1/K_d$) and n is the total number of cooperative sites.

Positively cooperative systems are easy to recognize in plots of f_R vs. [L] because of the characteristic sigmoidal shape of the curves. Consider the case of 1 ligand that binds to a receptor having two positively cooperative sites, where K are the association constants.

1 ligand, 2 sites, homocooperative:

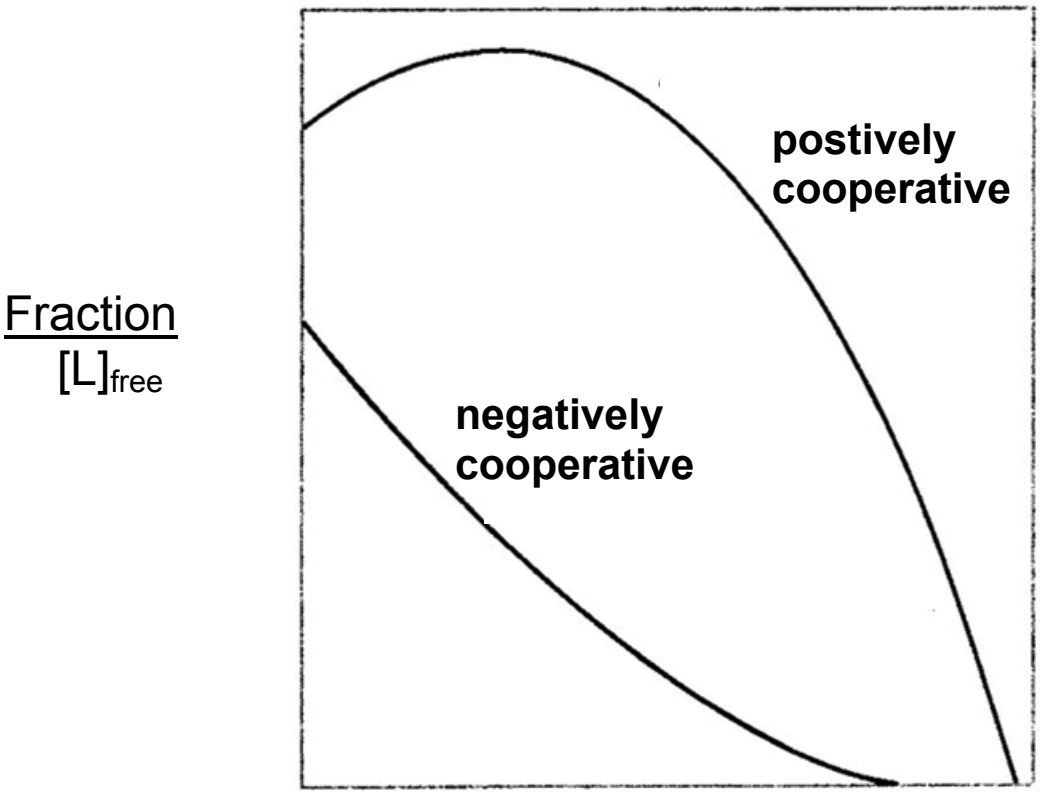
$$\mathbf{R + L \underset{K_1}{\rightleftharpoons} RL \underset{K_2}{\rightleftharpoons} RLL}$$

$$\mathbf{f_R = \frac{K_1[L] + K_1K_2[L]^2}{1 + K_1[L] + K_1K_2[L]^2}}$$



(Where f_R varies from 0 to 1.0) The relatively simple two site case is the simplest route to a sigmoidal binding curve, but there is an abundance of more complicated cooperative systems that can also lead to sigmoidal isotherms.

Cooperativity also leads to curved Scatchard plots (see below; recall that an alternate form of the Scatchard is to plot $[L]_{bound}/[L]_{free}$ vs $[L]_{bound}$).



We see that even for the very simplest case of homocooperativity, we get a very complicated equation (above). In practice, such equations are rarely used (think how many unknowns there will be to fit for!). What is sometime instead used is an all-purpose set of equations based on the Hill equation, an empirical equation.

For a generalized cooperative process going on involving n different sites, the following equation holds:

$$\text{fraction} = \frac{[L]_{free}^{\alpha}}{K_{d,app} + [L]^{\alpha}} \qquad (19)$$

In this case the maximum fraction of sites filled is always set to 1. This equation is known as the Hill equation. The K_d of this equation is not the true K_d, but is instead apparent (which makes sense, since in a cooperative process there will be *multiple* true K_d).

"α" is the Hill coefficient and is the "index of cooperativity". When α is determined to be equal to the number of binding sites, then it means the receptor is "maximally cooperative" (positive), meaning that once a single ligand binds, the affinity of the other sites for L become very high and they will be immediately filled. If α is 1 then there is no cooperativity. An α between 1 and its maximal value means that the degree of positive cooperativity is somewhere in between. Negative cooperativity shows up as $\alpha < 1$ with 1/(number of binding sites) being the lowest possible value (fully negatively cooperative).

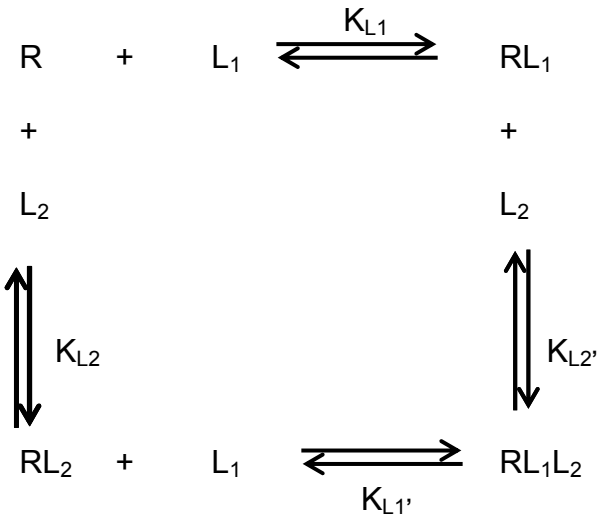
The Hill equation can be rearranged into a linar equation:

$$\log\left(\frac{\text{fraction}}{1-\text{fraction}}\right) = \alpha \cdot \log[L] - \log K_{d,app} \qquad (20)$$

This equation leads to the "Hill Plot", which provides a convenient way of determining α. $\log\{f_R/(1-f_R)\}$ is plotted vs. log[L]. The slope is the Hill coefficient and the y intercept = $-\text{Log}K_{d,app}$.

The Case of Heterocooperativity

What if there two different types of ligands (L_1 and L_2), both of which bind to different sites on the receptor, each with its own dissociation constant? If the sites are independent then $K_{d,1}$ and $K_{d,2}$ will be constant regardless of whether the site for the other ligand is occupied or not. However, what if the binding of one ligand to its own site perturbs the subsequent binding of the other ligand to its own site?:



In this case, one would likely study the fraction of binding for one site as a function of the concentration of ligand for that site and the concentration of the other ligand:

fraction of L_1 sites filled = function($[L_1]$,$[L_2]$)

What is the appropriate equation? If there is 1 site each for L_1 and L_2 and the fraction of L_1 sites complexed is monitored:

$$f_{R,L1} = [L_1] / \{[L_1]+(K_{L1}(1+[L_2]/K_{L2})/(1+[L_2](1-\beta)/K_{L2})\} \quad \textit{(needs checked)} \qquad (21)$$

where $\beta = (1-K_{L1})/K_{L2}$' (all dissociation constants). The 4th dissociation constant K_{L1}' does not appear in the above equation because it can be expressed as a function of the other three dissociation constants. The above equation is complicated!

The Case where Two Ligands Bind to the Same Binding Site (Competitively)

Even for the simplest case of a single receptor with a single binding site, there are several variables that must be considered:

1. Are the two ligands structurally identical, except that one is radiolabeled and one is not?
2. Will the experiment be run such that one can determine the fraction of receptor sites filled by each ligands, or only the total fractions of sites filled by both?
3. Will the fraction of L_1 and/or L_2 complexed be monitored (instead of the fraction of receptor sites filled)?
4. Is $K_{d,1}$ or $K_{d,2}$ known ahead of time?

If the two ligands are structurally non-equivalent the fraction of R that is complexed with L_1 is given by:

$$f_R = [L_1] / ([L_1]+K_{L1}(1+[L_2]/K_{L2})) \qquad (22)$$

where the K are dissociation constants and the concentrations are the free ligand concentrations.

When one of the competing ligands is the "natural" ligand for the binding site, the other is thought of as an "inhibitor", the "IC_{50}" gives the inhibitor concentration at which the binding of the "natural ligand" is inhibited by 50%. Unlike K_d, the "IC_{50}" has no standard state, but depends both on the K_ds of both ligands, and also upon the actual solution conditions in which it is measured. Hence IC_{50} tends to be rather subjective since it will be different (for a given pair of ligands) under differing conditions. Nevertheless, IC_{50} is frequently encountered in the literature. If K_d for the "inhibitor" is dubbed "K_i" and K_d for the "natural ligand" is simply "K_d", the following relationship holds:

$$K_i = \frac{IC_{50} \cdot K_d}{[L] + K_d} \qquad (23)$$

If the two ligands are chemically the same, but one of them (L') is radiolabeled the fraction of radiolabeled ligand that is bound will typically be measured as follows:

$$[L'R]/[R]_{total} = \frac{[L']}{K_d + [L] + [L']} \qquad (24)$$

where K_d is the dissociation constant (which is, of course, identical for both labeled and unlabeled ligand).

(For more detailed theory of competition assays, see Lepre et al., Chem. Rev. 104, 3641-3675 (2004)

A Survey of Methods Used to Determine the Fraction of Sites Filled or the Fraction of Ligand Complexed.

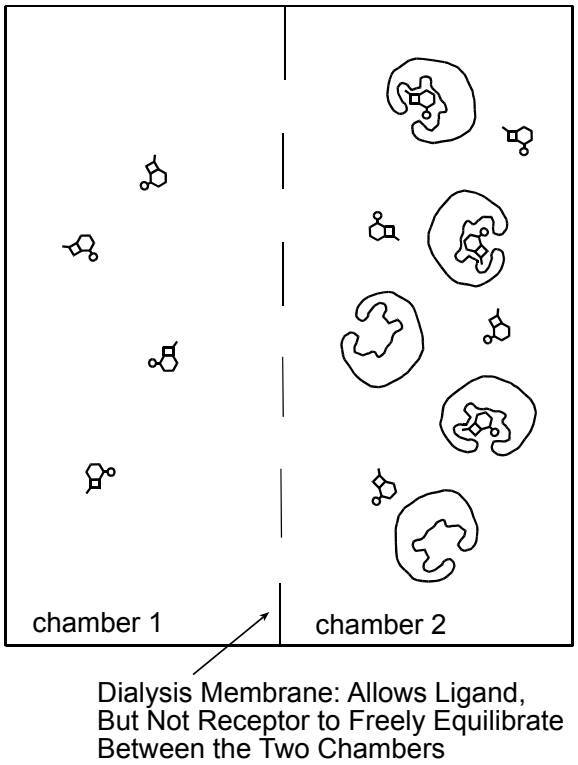
Equilibrium Dialysis.

Equilibrium dialysis involves putting the receptor (if it is a big molecule) into a dialysis tube or chamber that is separated from the surrounding reservoir solution by a semi-permeable membrane that will not allow the receptor to escape, but that will allow the free ligand (which must be much smaller) to freely diffuse into and out of the tube/chamber. The ligand is added to the outside solution and the system is allowed to equilibrate. When equilibrium has been reached (usually at least several hours) the total concentrations of ligand and (usually) the receptor inside of the bag/chamber are measured. Also the total concentration of the ligand in the outside reservoir solution is determined. This yields the following data:

$[R]_{total,\ outside} = 0$
$[L]_{total,\ outside} = [L]_{free}$
$[R]_{total,\ inside} = [R]_{free} + [RL]$
$[L]_{total,\ inside} = [L]_{free} + [RL]$

[RL] is easily calculated from this data as $[RL] = [L]_{inside} - [L]_{outside}$ and $[R]_{free}$ is determined:

$$[R]_{free} = [R]_{inside} - [RL].$$



We see from this that it is possible to determine K_d from a single point since:

$$K_d = \frac{[R] \cdot [L]}{[RL]}$$

However, a single point determination cannot confirm that a simple 1:1 binding model actually applies and also has a lot of error compared to a result obtained from many measurements. Thus, one would normally vary the ligand concentration and make several measurements and fit the data either to eq. 6 or a closely related equation (such as the Scatchard equation).

What if $[L]_{in}$ and $[L]_{out}$ can be measured but not $[R]_{in}$? In this case a series of $[L]_{free}$ vs. f_L data points would have to be collected and the data fit by the appropriate model to yield both $[R]_t$ and K_d.

The simplest type of equilibrium dialysis simply involves a dialysis bag in a beaker of solution. However, many companies now sell special equilibrium dialysis systems where there are two small chambers separated by a single layer of tubing.

Spectroscopic Methods

Often the ligand and/or receptor interacts with electromagnetic radiation (such as light) so that they either perturb the radiation in a measurable manner or respond to the input radiation by emitting their own characteristic signals. Examples of this include:

Ultraviolet Spectrophotometry: Light is passed through a solution and it is determined how much gets absorbed at one or more specific wavelengths.

Fluorescence Spectrophotometry: Shine light of a certain wavelength through the molecule and monitor, at one or more other wavelengths, light emitted by the molecule. If a range of wavelengths is scanned for emission, the plot of intensity vs. emission wavelength is known as an emission spectrum. Alternately, it is possible to vary the wavelength of the excitation light and measure intensity at a single emission wavelength. The plot of intensity versus excitation wavelength is known as an excitation plot.

Nuclear Magnetic Resonance Spectroscopy: In a strong magnetic field, the molecule is pulsed with broad frequency radio waves and the amount of radiation absorbed as a function of frequency is recorded. Each NMR-active atom will absorb to produce a signal at a discrete frequency.

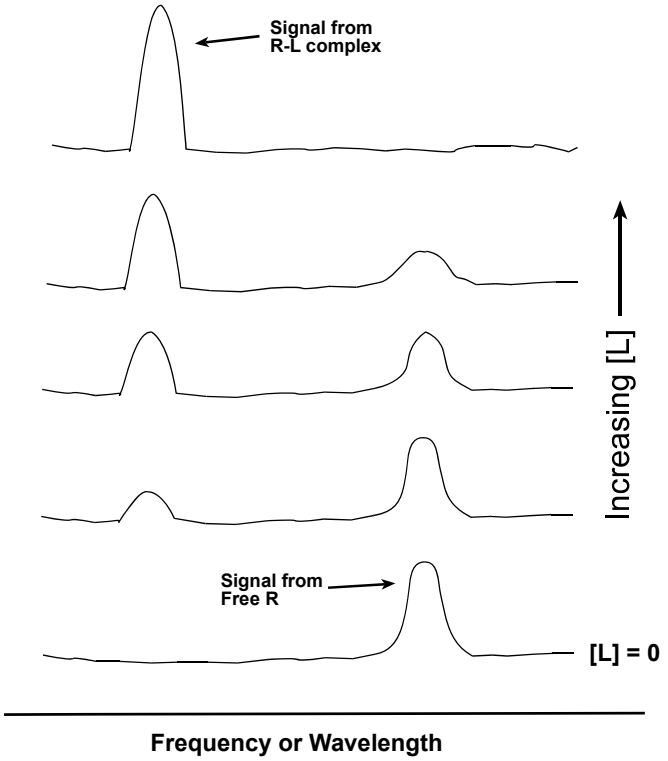
Circular Dichroism: Polarized light of a particular wavelength passes through solution followed by detection the difference in absorbencies of the right and left-handed light (the ellipticity, θ). The ellipticity is plotted as a function of wavelength.

For each type of spectroscopy the general idea is that a given free ligand and/or uncomplexed receptor has a characteristic "signal" or "signature" However, when a complex is formed, the spectroscopic properties of the molecule(s) as part of the complex may be very different than when free. In such cases, the changes in spectroscopic signals can be used to determine the relative concentrations of free and bound ligand and/or receptor: data that can be used to extract K_d.

For the moment let us assume that only the ligand (in free and complexed form) gives off a signal while the receptor is "spectroscopically silent". In this case, when there is a mixture of complexed and uncomplexed ligand, the observed spectra will fall into one of the following categories:

Binding/Dissociation is Slow on the Spectroscopic Time Scale:

Signals from both free and bound receptors and/or ligands are directly observed (sometimes getting stacked on top of each other). When the spectroscopic technique has a "shutter speed" is rapid relative to the rate of exchange, the technique will be able to resolve spectroscopic "pictures" of both the bound and free species, not a time average. In this case, the relative intensities (integrals) of each signal will be proportional to the relative concentrations (or "populations") of the molecule and the complex.



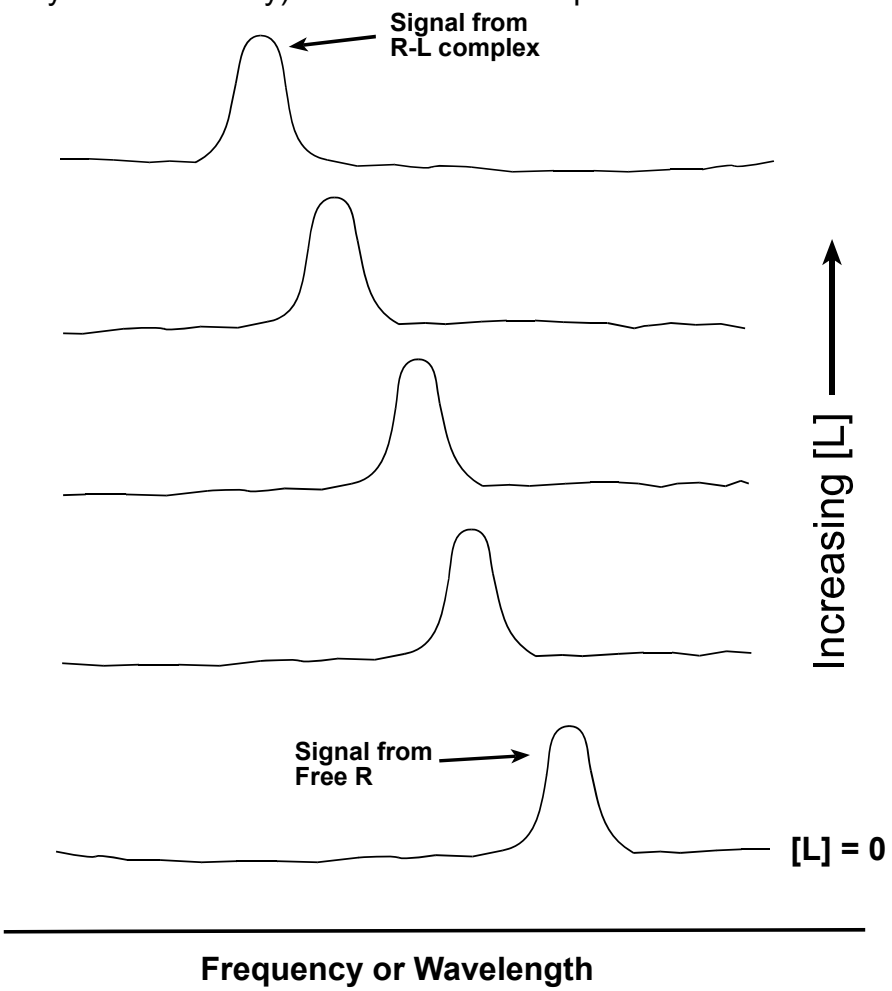
In this case, the relative intensity of the signal from the free ligand will be proportional to the free ligand concentration. The relative intensity of the signal from the complexed ligand will be proportional to the concentration of the bound ligand.

UV, fluorescence and CD will usually fall in this "rapid spectroscopy" or "fast shutter speed" category. Obviously, the use of these techniques in binding studies requires that the spectra from the free and complexed molecules not be identical.

Receptor-Ligand Association-Dissociation is Fast on the Spectroscopic Time Scale:

Some spectroscopic techniques have much slower "shutter speeds" so that they cannot distinguish between free and bound species. In these cases the observed signal is the

population-weighted average of the two species (only a single average signal is observed). This is often (but by no means only) the case for NMR spectra.



It is easy to see how the fraction of ligand or protein complexed can be determined from such a series of spectra. The changes observed in the spectra are proportional to the amount complexed vs. the amount free, which is easily converted to the fraction of ligand or receptor complexed)

It must be re-emphasized that these spectroscopic methods only work if the "intrinsic" signals from the free and complexed molecule are significantly different. This is not always the case.

Enzyme Kinetics and Dose-Response Pharmacology

We saw above that there is often a spectroscopic response to complex formation and that such responses can be used to quantitate molecular association. Often, binding of a ligand to a receptor brings about some physiologically-relevant response.

$$L + R \rightleftharpoons LR \longrightarrow \text{physiological response directly proportional to [LR] concentration}$$

For example, substrate binding to an enzyme leads to a chemical reaction. By measuring the ligand concentration dependence of the reaction rate we can sometimes determine the binding constant. Consider the Michaelis-Menten equation:

$$v = \frac{V_{max} \cdot [S]}{K_m + [S]} \qquad (25)$$

where is v the rate (the observed "response"), V_{max} (the maximum possible response) is observed only when all sites on the enzyme are filled with substrate, [S] is the free substrate (ligand) concentration, and K_m is the Michaelis constant (which is either equivalent or closely related to K_d). It is easy to rearrange this equation to yield:

$$\frac{v}{V_{max}} = \frac{[S]}{K_m + [S]}$$

This equation is very similar to equation 6, with v/V_{max} being equal to the fraction complexed. This make sense: $v = V_{max}$ only when the enzyme is fully saturated with substrate. "Fully saturated" means all sites filled, which means $f_{R,max} = 1.0$.

If the "response" to binding is not a chemical reaction, but a physiological process (such as a change in blood pressure or death rate in a population) a similar equation holds (assuming many assumptions are true):

$$\frac{\text{degree of response}}{\text{max possible response}} = \frac{[L]_{free}}{K_d + [L]_{free}} \qquad (26)$$

(Keep in mind that max possible response is observed only when all sites are filled). Because this type of assay is frequently used to test potential drug molecules, such binding studies are often called "dose-response" studies.

It should be noted that these equations only hold true in cases where the on/off rates are rapid compared to the "response" rate. This is known as the "rapid equilibrium" assumption. If this assumption is not valid (as is sometimes the case), then the above equations still hold, except that K_d becomes an *apparent* equilibrium constant- not the *true* dissociation constant. This apparent equilibrium constant is often referred to as the "IC_{50}": the "inhibitor" concentration at which 50% of maximal response is elicited.

Calorimetry

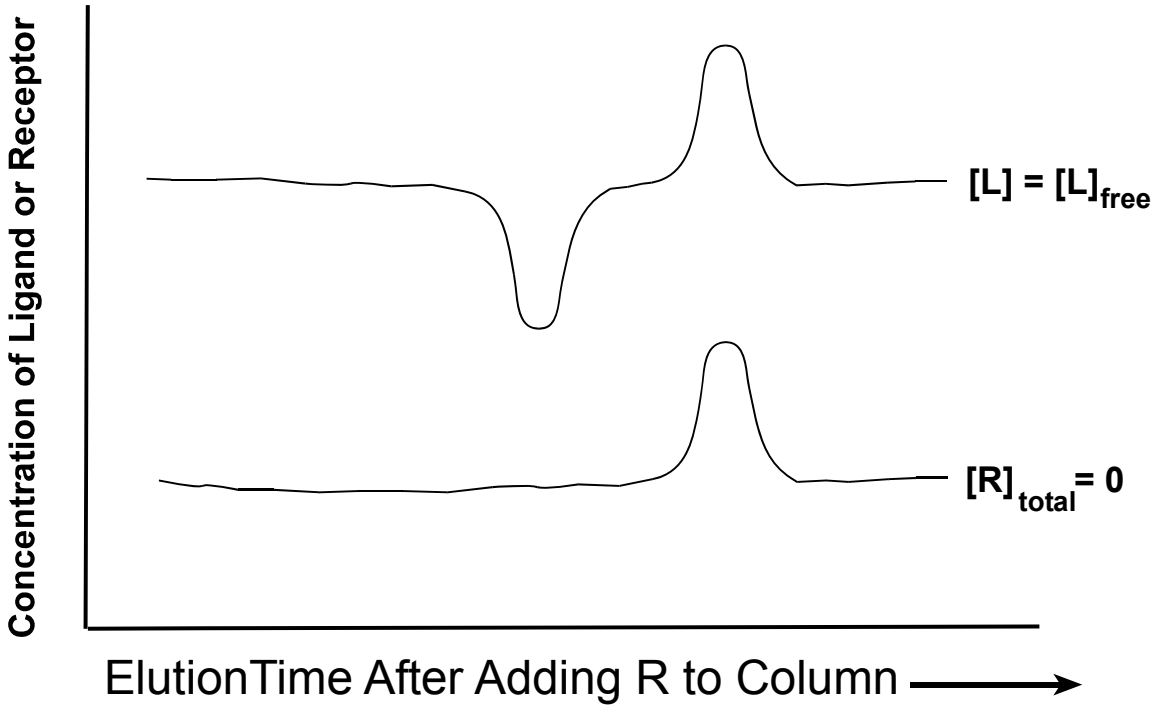
When a ligand binds to a receptor, heat will be generated (or absorbed). In calorimetry, a receptor is titrated with a ligand in a calorimeter and heat production is measured at each point in the titration. The heat generated is proportional to the amount of additional binding that takes place at that titration point. Binding isotherms are generated by plotting heat production vs. free [L] concentration. It is important to also account for heat that is produced during titration by secondary processes that have nothing to do with the binding event of

interest, such as heat generated by mixing stock ligand solution with the receptor solution and any non-specific binding of the ligand to the receptor or to other molecules/assemblies.

Chromatographic Binding Methods

There are at least two classes of such methods:

1. Gel chromatography-based methods. If one equilibrates a size exclusion chromatography column with an eluting buffer containing a particular $[L]_t$ concentration, a constant concentration of ligand will be flowing through and eluting from the column. Now, if a small volume of large molecule receptor-containing solution is applied to the column, followed by a resumption of elution with a constant [L] solution, the "pulse" of receptor in the column will immediately bind the ligand out of the nearby elution solution. Because the receptor (being a large molecule) passes through the column more quickly than the small molecules in the eluting buffer, it takes the bound ligand through more quickly, resulting in a surge of total ligand concentration coming off of the column (= bound + free ligand concentration) followed by a droop in the ligand concentration (because of depletion by the faster moving receptor), and then by a resumption of the normal ligand concentration. It is possible to determine the binding constant from a plot of [L] eluted from the column vs. time, if total [R] is known.



2. Affinity chromatography-based methods. In these methods a column is used with a chromatography column resin coated with tightly bound ligand or receptor. A solution of the receptor or ligand is applied to the column so that it gets stuck to the molecules on the resin. Then, elution is carried out by applying a low to high concentration gradient of a second molecule, whose affinity for the ligand or receptor is known. The concentration at which the molecule of interest displaces the receptor or ligand from the column is directly related to the dissociation constant. Note that this method requires that ligand/receptor complexes have

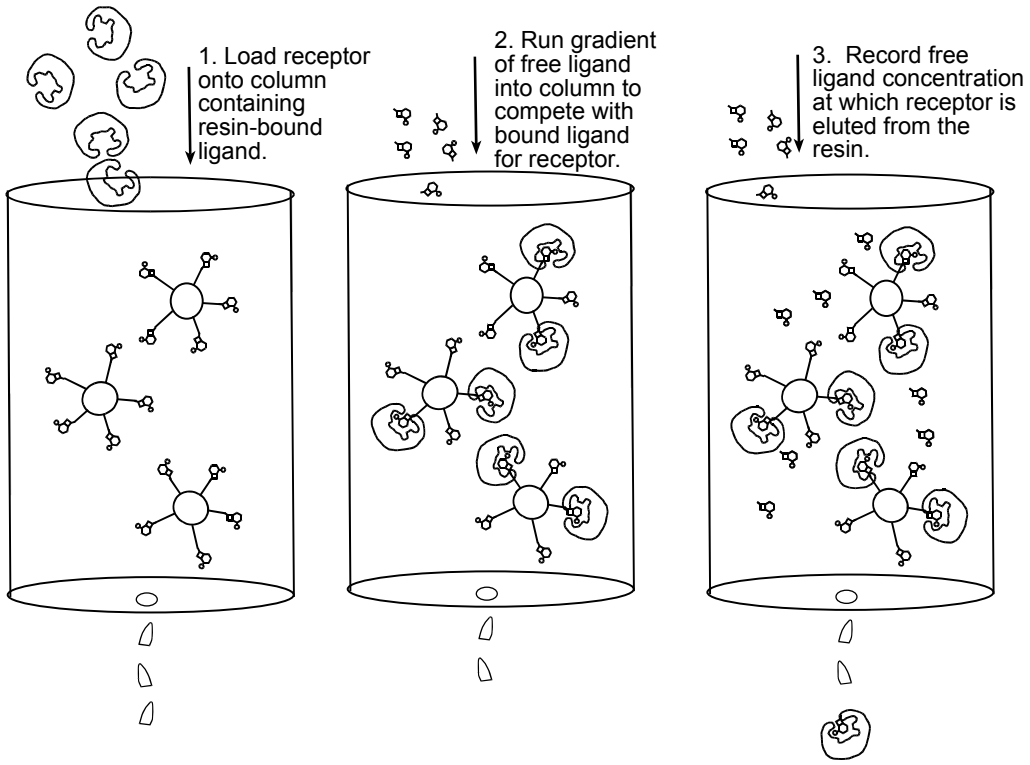
very long lifetimes: long on the time scale that it takes to run the column. Otherwise the bound receptor or ligand would be washed off of the column even the absence of a competitor. From the table above, it can be seen that K_d in the range of 100 nanomolar or tighter are required for the lifetimes of the complexes to be long enough for this approach to work.

Filter-Based Binding Assays

These methods are commonly used in pharmacology to detect binding of radioisotopically labeled ligand to membrane receptors. Usually, isolated cell membranes (that have both the receptor and many other membrane proteins), whole cells, or lysed cells are used. In other words, these experiments typically involve looking at interactions between a small molecule and a receptor that is part of a gigantic (filterable) superstructure.

The cells/membranes containing the receptor are mixed with the ligand. The mixture is then filtered so that the cells/membrane containing the receptor and bound ligand are stuck to the filter. The filter is then washed to remove any residual unbound ligand and then subjected to scintillation counting to detect and quantitate the amount of bound ligand. An important control experiment is to repeat measurement on cells/membranes that are identical except that they have no receptor. By repeating measurements under conditions where a cold (non-labeled) ligand is used to compete off the hot ligand, it is possible to get binding curves.

This is another example of a method that works well only if the complex between the receptor and the ligand is long-lived relative to the time it takes to carry out filtration and then wash off excess unbound ligand.

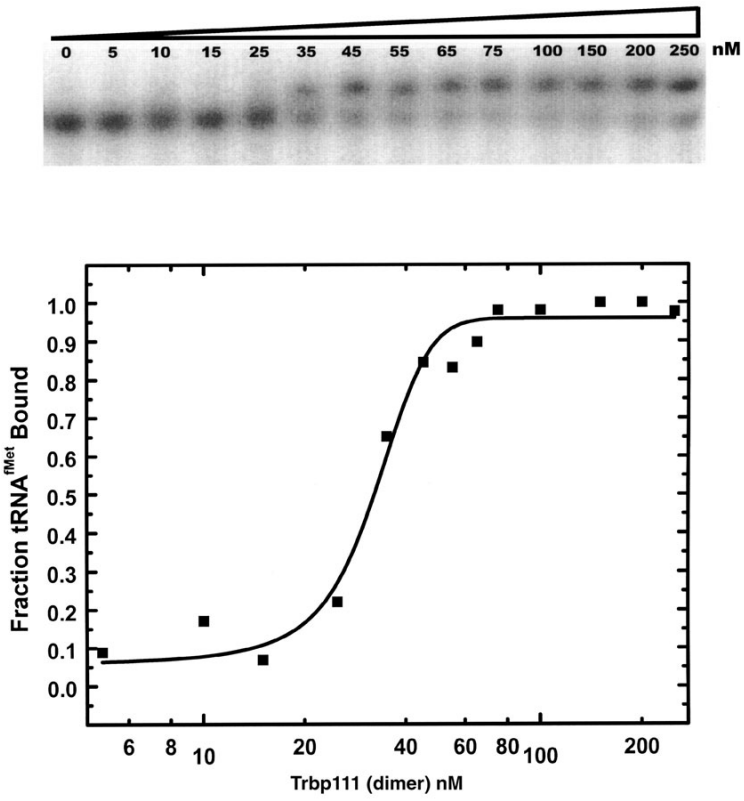


Gel Mobility-Shift Binding Assay

In polyacrylamide gel electrophoresis a polymer matrix is used, which is a network of interconnected pores of fairly uniform size. The size of the pores is determined by the % cross-linking of the polymer-- higher cross-linking means smaller pores. To a well in one edge of a slab of such a polymer is loaded protein and/or DNA. An electric field is then applied across the plane of the slab. The molecules in the sample well are then electrically dragged into the gel by the current and will migrate through the gel. Other factors being equivalent, large molecules tend to migrate slower because it is harder for them to get through the pores.

What if both DNA/RNA and a protein that binds to the nucleic acid is loaded into the sample loading well of the gel? Once equilibrium has been reached, there will be both free and complexed protein/nucleic acid populations in the well. When the current is turned on everything is quickly pushed into the gel where the free nucleic acid or protein (whichever is lightest) is quickly separated from the complex (because of very different molecular weights). As the gel runs, very little of the complex dissociates because it is confined in the pores of the gel.

Thus, after running the gel, the exact amount of free nucleic acid or protein in the original sample (and sometimes the concentrations of the complexes) can be easily determined by autoradiography, densitometry, or scintillation counting. By running several lanes in which the amount of ligand is varied (while the receptor is held constant), it is possible to obtain the ususal fraction$_R$ as a function of [L]$_{free}$ data, from which a dissociation constant can be determined. Typically, the oligonucleotide serves as the "receptor" in these studies and is radioactively labeled, while the concentration of the unlabeled protein ligand is varied from lane to lane.



Binding of tRNA$_f^{Met}$ to Trbp111. (A) Gel-motility shift assay of Trbp111 binding to radioactively labeled tRNA$_f^{Met}$. Trbp111 was incubated at increasing concentrations with 1−2 nM [5'-^{32}P] tRNA$_f^{Met}$ and electrophoresed on a native acrylamide gel at 4°C. The top band denotes formation of a complex. (B) Binding profile obtained after determination of fraction of tRNA$_f^{Met}$ bound to Trbp111 (squares) at various concentrations. Data were simulated using the curve corresponding to a simple-binding equilibrium (solid line). The apparent dissociation constant for tRNA$_f^{Met}$ binding was determined as 32 nM (dimer concentration). (From: Morales et al., *The EMBO Journal* (1999) 18, 3475–3483.)

One limiting factor for these experiments seems to be the "dead time" it takes the DNA and protein to enter at the beginning after turning the current on. If all of the sample does not enter the gel very rapidly, changes in the concentrations in the loading wells can cause problems. Since the time it takes all sample to enter into gel is on the order of seconds, the technique may tend to be limited to fairly tight (or slow binding) where the $t_{1/2,off}$ is also on the order of seconds or longer. This condition that is frequently met in the case of protein-nucleic acid interactions, for which binding is often avid.

Ligand Competition Assays

Frequently, more than one type of ligand can compete for the same binding site. In such cases the binding of "Ligand 1" to the receptor can be studied as a function of the amount of "Ligand 2" bound to the receptor.

$$L_1\text{-}R \rightleftharpoons L_1 + R + L_2 \rightleftharpoons L_2\text{-}R$$

We already covered the theory behind this class of experiments. There are two forms of these experiments to be especially aware of are:

1. L_1 is exactly the same as L_2, except that one of them is radiolabeled. A typical study involves forming a complex of labeled ligand with the receptor and then adding aliquots of unlabeled ligand and watching the degree of displacement of the labeled ligand from the complex.

2. The case where the dissociation constant of Ligand 1 is known in advance. This makes determination of K_d for L_2 very easy.

A Checklist of System Variables to Consider When Planning Binding Studies

Is the receptor pure? If it is not pure, are there other receptors or sites to which the ligand could also bind? If so, then the total ligand concentration will be:

$$[L]_t = [L]_{free} + [L]_{bound\ to\ receptor} + [L]_{bound\ to\ other\ molecules}$$

This may be a serious complication if the third term is similar in size to either the free or receptor-associated ligand concentrations. If the receptor is not pure, it may be difficult to quantitate exactly how much of it is there- this can complicate analysis.

Is the ligand pure? If not, there will similar complications as for impure receptor. Can "impurities" also bind to the receptor and compete for the same binding sites as the ligand of interest? Do impurities make it hard to determine the ligand concentration?

Are the receptors and ligands stable during the time it takes to do a binding study? If not, how does one deal with the fact that the total concentrations of ligand and/or receptor are degrading during the course of an experiment?

Is receptor or ligand associated with membranes? micelles? If the receptor is membrane bound then one may have to be concerned that the ligand may bind to the membrane as well

as directly to the receptor site. Is this taken into account by the model? If "reconstituted" model membranes like liposomes or micelles are used, does the reconstitution procedure perturb the binding properties of the receptor from its properties in the original native membranes? If it is inserted into liposomes, is the "business end" of the receptor located on the outer side of the bilayer? on the inside? on both sides?

If in vivo studies are being carried out, can the ligand physically make it to the receptor? Lots of things can happen in a real physiological system to a ligand. For example, it may be chemically degraded or altered by an enzyme. It may not be able to reach the target receptor of interest because it cannot get into the cell or because, once in the cell, a transport protein may kick it out again.

If the receptor is an enzyme, is the ligand a substrate? If so then it may be possible be able to study binding using enzyme reaction kinetic methods, but not using equilibrium methods (unless some enzyme cofactors or co-substrates are missing), because the ligand will be converted into product.

Will the ligand bind metal ions or other small molecules present in the system? If so, how tight is the binding? Will multiple types of metal-ligand complexes be present? Which forms of these complexes will bind to the receptor and how does one sort out all of the possibilities?

If the ligand has ionizable moieties, what will be the ionic form under the conditions of the experiments? Will multiple ionic forms be present? Which ionic forms of the ligand will bind to the receptor and how will the possibilities be sorted out?

If one is studying a binding process in vitro, can results be extrapolated to the natural in vivo process?

If binding is being monitored as a physiological response to the ligand, is it certain that the response is the *direct* result of ligand-receptor association, or could it be an indirect connection? Maybe the ligand activates a protein that stimulates a protein kinase that phosphorylates a receptor that is then activated to bring about the physiological response of interest. If so, is the response really proportional to the original extent of the association of the ligand with receptor?

A Checklist of Experimental Variables to Consider in Planning Binding Studies

How tight is binding likely to be and will the chosen technique be well matched to determine binding constants in the expected range? For example, NMR is good at measuring relatively weak binding constants (K_d > 100 μM), but not effective for tighter binding. Conversely, techniques that rely on physically separating free ligand from bound ligand often will only work if binding if very tight (long-lived complexes).

What is the time resolution of the chosen technique? In other words, does the chosen method for quantitating the concentrations of free ligand/receptor and/or the complex measure the time average concentrations of a system that is at equilibrium during the time it takes to make a measurement, or can the chosen method be used in a time resolved manner to follow the approach of a mixture of ligand and receptor to reach equilibrium. If

spectroscopy is being used to monitor binding, does the technique allow one to directly observe and quantitate both free and complexed species, or does the technique give a signal that reflects a population-weighted average between free and bound species?

Will binding be monitored under equilibrium conditions? If not, binding may best be characterized using kinetic methods (measuring k_{on} and k_{off}), which may require very different models for data interpretation.

Will the extent of binding be monitored directly or will it be monitored indirectly through a process such as enzyme catalysis or a physiological event that is thought to be coupled to the binding of the ligand and receptor of interest? In these cases, be aware of all of the assumptions required by such methods if the interpretation of the data in terms of binding is to be valid.

How direct is the relationship between the measured parameter in a binding study and the degree of binding? (See below)

Can the total, free, and complexed concentrations of all of the relevant ligands and receptors present in the system be directly quantitated? As noted earlier, if the answer is no, it may complicate analysis.

Will the total ligand concentration be large (at all points in the experiment) compared to the concentrations of the Ligand-receptor complex? If so, then data interpretation may be simplified.

Can control experiments be designed that will eliminate all possible ambiguity in results?

Is the system stable enough and available in sufficient quanitity to take multiple points?

What is known about the system in advance? Is the stoichiometry known? Is it know whether there is any cooperative binding? etc.

There are a lot of things to think through before (and after) embarking upon a binding study!

Acid-Base Equilibria

A "Bronsted acid" is a molecule that can give up a H^+, while a "Bronsted base" can accept one. In biochemistry this usually means giving up or accepting protons in aqueous (water) solutions:

$$AH + H_2O \rightleftharpoons A^- + H_3O^+$$

When we refer to H+ as acid, we really mean H_3O^+. For the common definition of pH: $pH = \log(1/[H^+])$, it is more accurgate to define: $pH = \log(1/[H_3O^+])$. In this section H^+ and H_3O^+ will be used interchangeably.

In the above equilibrium, we there is a great fundamental paradox: "A" exists in two forms- in one form it is clearly an acid and in the other form it is a base! This is true of most acids and

bases and for this reason molecular pairs represented by AH and A^- are said to be related as the "conjugate acid" and "conjugate base" of one another.

Water is kind of unique because when it dissociates it forms both acid and base:

$$2\ H_2O \rightleftharpoons H_3O^+ + OH^-$$

This is like any other equilibrium. For pure water:

$$K_{eq} = \frac{[H_3O^+]\,[OH^-]}{[H2O]_2} = \frac{[1 \times 10^{-7}]\,[1 \times 10^{-7}]}{[55]^2} = 3.3 \times 10^{-18}$$

Since $[H_2O]$ will always be 55 M, it is traditional to drop the bottom of this equation and redefine K_{eq} as "K_w", where:

$$K_w = [H_3O^+][OH^-] = 1 \times 10^{-14}$$

From the above, we see that the concentrations of free H^+ and OH^- in neutral water are very low. Water is a very "weak" acid and base.

This leads to a definition. The pH of an aqueous solution is defined as:

$$pH = -\log[H^+] = \log(1/[H^+])$$

while the pOH is defined as

$$pOH = -\log[OH^-] = \log(1/[OH^-]).$$

Thus, for pure water pH = pOH = 7.0.

pH and pOH are related:

$$pH + pOH = 14 \qquad \text{(always)}$$

Strong Acids and Bases

HCl in water dissociates in water, effectively to infinity:

$$HCl + H_2O \longrightarrow H_3O^+ + Cl^-$$

From this equation, it is seen that Cl^- is the "conjugate base" of HCl. However, in aqueous solution, this equilibrium lies very far to the right because Cl^- has virtually no tendency to pick up a proton (if it did, NaCl would be a base!). Hence, while Cl^- may technically be a conjugate base, in practice it isn't. For this reason, HCl and certain other acids such as H_2SO_4 and HNO_3 are referred to as "strong" acids, meaning they can be assumed to totally dissociate when dissolved in aqueous solution.

Acids and bases are usually supplied as highly concentrated aqueous solutions. The actual molar concentrations can be obtained from various sources (such as a table that appears in the Merck Index). Concentrated HCl is 11.6 M in HCl (it is prepared by bubbling HCl gas through water until the water is saturated with the acid).

Just as there are strong acids, there are strong bases as well- most notably the alkali metal hydroxides (NaOH, KOH). Be aware that while we think of ammonia (NH_4OH) as a strong base from a household standpoint, it is not a strong base from a chemical standpoint: it does not fully dissociate into NH_4^+ and OH^- in water.

When a certain amount of a strong acid or a strong base is added to water, it is very easy to calculate the pH. Since the equilibrium constant is effectively infinity in favor of H^+ or OH^- formation, the moles of strong acid or strong base added is equal to the moles of H^+ or OH^- that form in solution. For example, what if 1 ml of concentrated HCl is added to 1 liter of water?

final $[H^+]$ = moles HCl added/liter final solution + initial conc. of H^+ in water

= (0.001 liters conc. HCl · 11.6 moles HCl/liter conc. HCl
· 1/(1 liter final solution) + 1 X 10^{-7} M =

0.01160001 M = $[H^+]$

So the pH is log(0.0116)-1 = 1.9

Weak Acids and Bases

From the above we saw that in the case of a strong acid or a strong base, the assumption is made that dissociation in water is infinite. This is obviously not the case for water, nor for other weak acids and bases. For example, a little acetic acid is added to water, the pH goes down, but not nearly as much as for an equimolar amount of a strong acid: this is because some protons remain attached to the parent acid. The amount by which the pH goes down is determined by the degree of dissociation:

$$HOAc + H_2O \rightleftharpoons H3O^+ + OAc^-$$

The position of this equilibrium for HOAc dissolved in pure water is determined (big surprise) by the dissociation constant. However, by convention, for acids and bases the H_2O concentration is not included in acid base equilibrium constants. Thus:

$$K_d = \frac{[H^+]\,[OAc^-]}{[HOAc]}$$

This dissociation constant is typically referred to as the acid dissociation constant (K_a, not to be confused with the association constant). Similarly, if a conjugate base form of acetic acid, sodium acetate, is dissolved in water, the acetate ion could accept a H^+ from water:

$$OAc^- + H_2O \rightleftharpoons HOAc + OH^-$$

and the "base dissociation constant': could be determined:

$$K_b = \frac{[HOAc]\,[OH^-]}{[OAc^-]}$$

K_{acid} and K_{base} for all conjugate acid/base pairs are related:

$$K_a \cdot K_b = 1 \cdot 10^{-14}$$

So, acids with relatively large K_as are relatively strong acids (they tend to release their protons into solution) and a similar relationship hold for K_bs and bases.

What all does K_a (or K_b) tell us? As we learned earlier, for any binding process the dissociation constant provides (among other things) the ligand concentration at which the "receptor" is half saturated by a "ligand" (half of the sites are filled). In this case H^+ can be thought of as "the ligand" and the conjugate base that gets protonated can be thought of as "the receptor". Thus, if we know the K_a, we know the $[H^+]$ at which the "receptor" is 50% protonated (half of the molecules will be in acid form, half in base form).

Secondly, if K_a (or K_b) is known for an acid and we throw some into water, the $[H^+]$ concentration (and hence the pH of the solution) can easily be calculated.

Just as we don't talk about $[H^+]$ in a real lab situation (we use "pH"), we don't talk much about K_a and K_b. Instead, we usually talk about pK_a (equal to $-\log K_a$) or pK_b (equal to $-\log K_b$). This is useful, because the pK_a will tell us the pH at which the molecule will be 50% acid form, 50% base form. The significance of this becomes apparent when we talk about buffers below. It should go without saying that the pK_a and pK_b for a conjugate acid/base pair are related:

$$pK_a + pK_b = 14$$

What determines the pK_a of a substance? One factor is the nature of the chemical moiety and its covalent environment. For example, amines are usually bases- becoming protonated in the pH 8-12 range, carboxylic acids are usually acidic having pK_as in the pH 2-5 range. Secondly, the local environment of the molecule can influence pK_a. High salt concentrations typically reduce the pK_a (because the salt cations compete with protons for the lone pairs). Low polarity environments can shift a pK_a in either direction, favoring either the acid or base form depending upon which is neutral.

Buffers

In the above section we learned that weak acids/bases exist in both conjugate forms in aqueous solutions. Let's think about this a bit more. Acetic acid has a pK_a of 4.8 (hence a K_a of 1.4×10^{-5}). What will the pH be if 10 ml of glacial acetic acid (17.4 M) are dissolved into a liter of water?

$$pH = \log(1/[H^+])$$

$[H^+] = 1 \times 10^{-7}$ (from neutral H_2O) + the $[H^+]$ from acetic acid dissociation

$[H+]_{\text{from acetic acid}} = ???$

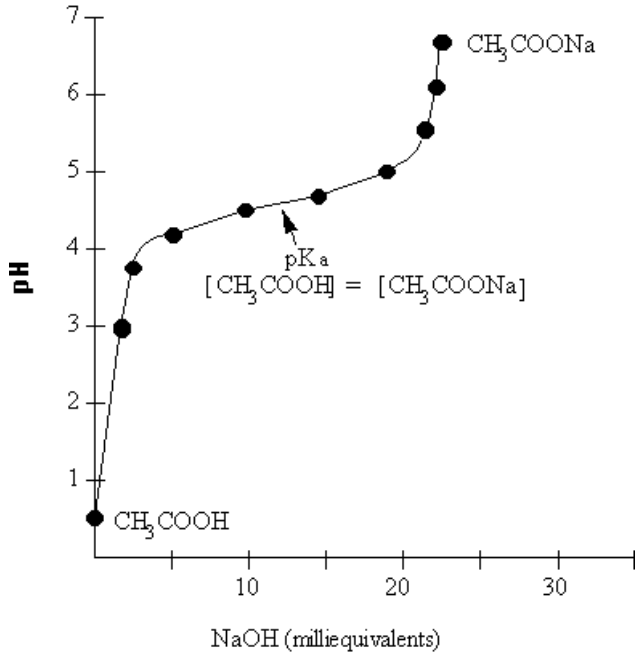
[Acetic acid total] = 0.010 liters · 17.4 moles/l · 1 liter = 0.17 moles/liter = [acid form] + [acetate ion]

So...

$$K_a = 1.4 \times 10^{-5} = \frac{[x] \cdot [x]}{[0.17 - x]} \qquad \text{where } [x] = [OAc^-] = [H^+]$$

This can be solved for x using the quadratic equation leading to $[H^+]$ = 0.0015 M and pH = 2.8. As would be expected if only the acid form is dissolved in water, the pH ends up below the pK_a for the acid.

For this 0.17 M acetic acid solution (where pH = 2.8), what happens if we now titrate in pure OH^- (in the form of a strong base like NaOH)? What is initially observed is that the pH changes only slowly with the amount of added base, until at some point it shoots up:



What is going on here? Remember, the acid is part of an equilibrium:

$$AH \rightleftharpoons A^- + H^+$$

with the "position" of the equilibrium being determined by K_a. Now, when OH^- is added (as NaOH or KOH), the situation becomes more complex:

$$AH \rightleftharpoons A^- + H^+ + OH^- \rightleftharpoons A^- + H_2O$$

(In the above, it is presumed that OH^- does not directly steal AH's proton, but only combines directly with H^+. This is really not true, but for the purposes of this discussion it is an acceptable working approximation.) As shown above, the second equilibrium lies almost infinitely in favor of OH^- combining with H^+ to form H_2O. Thus, after adding OH- overall equilibrium is restored only when all of the added OH^- has been neutralized by H^+. What is the effect of OH^- addition upon the first of the two equilibria? Well since the original equilibrium concentration of H^+ gets soaked up (in part) by the added OH^- more AH dissociates to restore equilibrium. Thus, even though OH^- has been added the ability of AH to dissociate dictates that the equilibrium H^+ concentration (and therefore pH) doesn't change all that much. In this manner AH acts as a buffer of the pH and can continue to do so (as more OH^- is added) until the added OH^- has depleted all of available protons (AH is all gone, only A^- remains) and the pH shoots up when there are no more AH left to donate a proton to pair with the added OH^-.

The above process illustrates how a weak acid or a weak base can act as a pH buffer. We saw that as long as there was enough AH around, the addition of base really didn't change the pH very much. This same thing would hold true if a strong acid is added to a solution of A^-. A^- would "soak up" the added acid until it was completely converted to AH, at which point further addition of H^+ would cause the pH to plummet.

From this, we can see that weak acids and weak bases can be used to buffer solutions pH-wise against dramatic variations in pH when small amounts of strong base or acid are added.

What if one needs to maintain a solution at a certain pH in a situation in which some strong acid or strong base might be added? In this case, it would not be appropriate to have a situation where the buffer was almost exclusively in either its acid or base forms. This is because each will only buffer against added base or acid, respectively. However, what if the solution contains a buffer that is already 50% acid/50% conjugate base at the pH one wishes to work at? In this case, the buffer could soak up both strong acid and strong base.

How does one find a buffer that will be the desired 50/50 conjugate acid/base at a given pH? This is easy. Find an acid/base whose pK_a is very near to the working pH: when pH = pK_a the buffer molecule population will be 50% in acid form and 50% in base form.

Where can lists of buffer/pK_as be found? Tables. (See Table below).

It should be noted that a number of compounds have more than 1 pK_a because they can lose more than 1 proton. For example, phosphoric acid:

$$H_3PO_4 \underset{K_{a,1}}{\rightleftharpoons} H_2PO_4^{1-} + H^+ \underset{K_{a,2}}{\rightleftharpoons} HPO_4^{2-} + H^+ \underset{K_{a,3}}{\rightleftharpoons} PO_4^{3-} + H^+$$

Choice of a Buffer

Calbiochem will send a nice booklet on buffers on request- see their catalog. Segel has an exhaustive listing of buffers and pK_as in its appendices.

The rule of thumb is that a buffer needs to have a pK_a that is within 1 unit of the working pH if the buffer is to be effective.

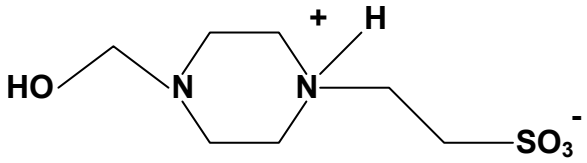
One should anticipate how much acid or base may be produced in the course of an experiment and make sure the buffer is suitable to maintain the pH. On the other hand, it is undesirable to use too high of a concentration (rarely > 0.1 M) because this can lead to various artifacts. In calculating the expected pH change during a process, the "Henderson-Hasselbach" equation can be very helpful:

$$pH = pK_a + \log([A^-]/[HA])$$

Zwitterionic buffers are preferred to anionic or cationic buffers. A zwitterion is a molecule that has both positive and negative charges, but in which the net charge is zero. There are a whole series of buffers known as the "Good" buffers that will be zwitterionic in at least one of their conjugate forms. These are listed in the Sigma catalog and include HEPES, PIPES, MOPS, etc.

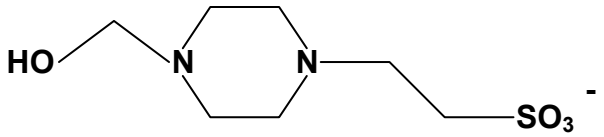
HEPES (pK_a = 7.55)

acid form (zwitterionic):



pH << 7.4

base form:



12 >> pH >> 7.4

The advantage of zwitterionic buffers over ionic buffers is that they are very non-reactive and are less likely to produce experimental artifacts. For example, phosphate (negatively charged) will bind to metal ions that can cause troubles in studies of the effect of metal ions on biologically relevant processes. Zwitterions, with their net 0 charges have less affinity for metal ions. The only problem with zwitterionic buffers is that they tend to be expensive. Thus, ionic buffers such as Tris (positively charged) and phosphate remain popular, especially for high volume procedures.

Tris buffers are not very compatible with standard Ag/AgCl pH electrodes (I think because Tris has an unusual avidity for silver). Thus, calomel (mercury-based) electrodes are usually used.

The pK_a of buffers and the pH of their solutions is temperature dependent (see Table, as high as 0.03 pH or pK_a units per degree of deviation from the 25º values). This should be kept in mind when performing studies on ice or in a warm water bath.

Borate buffers are not suitable for working with nucleic acids because it reacts with the diol of the ribose moieties to form insoluble complexes.

Very high salt concentrations can make the pH meter readings unreliable- in such cases one must dilute the solution by a factor of 10 or so and then measure the pH. Assuming the solution is buffered, the pH will change little due to dilution.

Some buffers will specifically interact with the biomolecules of interest. For example, one would not want to use phosphate buffer when studying phosphatases (enzymes that hydrolyze phosphoesters).

When working with solutions containing detergents (as when working with membrane proteins), buffers should be mixed and their pH should be adjusted BEFORE adding detergent, as some detergent types interfere with pH measurements.

The following table is taken from Calbiochem's booklet "Buffers".

Table 1: pK_a Values for Commonly Used Biological Buffers and Buffer Constituents

Product	Cat. No.	M.W.	pK_a at 20°C
ADA, Sodium Salt	114801	212.2	6.60
2-Amino-2-methyl-1,3-propanediol	164548	105.1	8.83
BES, ULTROL® Grade	391334	213.2	7.15
Bicine, ULTROL® Grade	391336	163.2	8.35
BIS-Tris, ULTROL® Grade	391335	209.2	6.50
BIS-Tris Propane, ULTROL® Grade	394111	282.4	6.80
Boric Acid, Molecular Biology Grade	203667	61.8	9.24
Cacodylic Acid	205541	214.0	6.27
CAPS, ULTROL® Grade	239782	221.3	10.40
CHES, ULTROL® Grade	239779	207.3	9.50
Citric Acid, Monohydrate, Molecular Biology Grade	231211	210.1	4.76
Glycine	3570	75.1	2.34[1]
Glycine, Molecular Biology Grade	357002	75.1	2.34[1]
Glycylglycine, Free Base	3630	132.1	8.40
HEPES, Free Acid, Molecular Biology Grade	391340	238.3	7.55
HEPES, Free Acid, ULTROL® Grade	391338	238.3	7.55
HEPES, Free Acid Solution	375368	238.3	7.55
HEPES, Sodium Salt, ULTROL® Grade	391333	260.3	7.55
HEPPS, ULTROL® Grade	391339	252.3	8.00
Imidazole, ULTROL® Grade	4015	68.1	7.00
MES, Free Acid, ULTROL® Grade	475893	195.2	6.15
MES, Sodium Salt, ULTROL® Grade	475894	217.2	6.15
MOPS, Free Acid, ULTROL® Grade	475898	209.3	7.20
MOPS, Sodium Salt, ULTROL® Grade	475899	231.2	7.20
PIPES, Free Acid, Molecular Biology Grade	528133	302.4	6.80
PIPES, Free Acid, ULTROL® Grade	528131	302.4	6.80
PIPES, Sodium Salt, ULTROL® Grade	528132	325.3	6.80
PIPPS	528315	330.4	3.73[2]
Potassium Phosphate, Dibasic, Trihydrate, Molecular Biology Grade	529567	228.2	7.21[3]
Potassium Phosphate, Monobasic	529565	136.1	7.21[3]
Potassium Phosphate, Monobasic, Molecular Biology Grade	529568	136.1	7.21[3]
Sodium Phosphate, Dibasic	567550	142.0	7.21[3]
Sodium Phosphate, Dibasic, Molecular Biology Grade	567547	142.0	7.21[3]
Sodium Phosphate, Monobasic	567545	120.0	7.21[3]
Sodium Phosphate, Monobasic, Monohydrate, Molecular Biology Grade	567549	138.0	7.21[3]
TAPS, ULTROL® Grade	394675	243.2	8.40
TES, Free Acid, ULTROL® Grade	39465	229.3	7.50
TES, Sodium Salt, ULTROL® Grade	394651	251.2	7.50
Tricine, ULTROL® Grade	39468	179.2	8.15
Triethanolamine, HCl	641752	185.7	7.66
Tris Base, Molecular Biology Grade	648310	121.1	8.30
Tris Base, ULTROL® Grade	648311	121.1	8.30
Tris, HCl, Molecular Biology Grade	648317	157.6	8.30
Tris, HCl, ULTROL® Grade	648313	157.6	8.30
Trisodium Citrate, Dihydrate	567444	294.1	–
Trisodium Citrate, Dihydrate, Molecular Biology Grade	567446	294.1	–

1. $pK_{a1} = 2.34$; $pK_{a2} = 9.60$
2. $pK_{a1} = 3.73$; $pK_{a2} = 7.96$ (100 mM aqueous solution, 25°C).
3. Phosphate buffers are normally prepared from a combination of the monobasic and dibasic salts, titrated against each other to the correct pH. Phosphoric acid has three pK_a values: $pK_{a1} = 2.12$; $pK_{a2} = 7.21$; $pK_{a3} = 12.32$

Acetic acid *pKa = 4.75* *Ammonia* *pKa = 9.3*
Formic acid *pKa = 3.75*
CAPS *pKa = 10.4*
Phosphoric Acid *pKa1= 2.1*

Buffers: A Summary

All buffers are *weak* acids/bases, meaning K_a is finite. For acetic acid in water:

$$H_3C\text{-}COOH + H_2O \underset{}{\overset{K_a}{\rightleftharpoons}} H_3O^+ + H_3C\text{-}COO^- \qquad pK_a = \log(1/K_a)$$

(acid form) (base form)

All weak acids have weak base "conjugate" forms and vice versa.

When the pH equals the pKa, the moiety will be 50% in its acid form and 50% in its base form.

When the pH is more than 1 unit lower than the pKa, then >90% will be in the acid form.

When the pH is more than 1 unit higher than the pKa then >90% will be in the base form.

Buffers are effective at controlling pH when:

(1) the concentration is high enough to "soak up" any acid or base produced in the solution

(2) the target pH is within 0.5-1 units of the buffer's pKa.

Chelating Agents

Often, it is desirable to work under conditions in which there are not divalent or trivalent cations present in solution. This is very easy to do by including a certain amount of a "chelating agent" in the buffer. Such agents bind to and thereby "tie up" all of the stray divalent or trivalent metal ions present. These molecules are kind of molecular octopuses that bind cations because of having several negatively charged carboxyl groups tethered closely but flexibly together and that "gang up" to bind metal ions:

What should one know about chelating agents?

1. Carboxylates are weak bases (the conjugate bases of carboxylic acids). Thus, when EDTA or EGTA are dissolved, they can perturb the pH of the solution (depending upon their original ionic form).

2. Once one or more of the carboxylates becomes protonated, the affinity of the chelator for metal ion is greatly reduced (because protons are successfully competing with the "ligand binding sites"). Therefore, be aware that the lower the pH, the harder it will be for a chelator to do its job.

3. It is important to have more moles of chelator in the solution than moles of divalent/trivalent metal ions or the chelator will become saturated with free ions left over.

4. Chelators have different affinities for different ions. For example, EGTA is has an especially high affinity for Ca^{2+}, but may be unsuitable for other ions. EDTA is used for most other divalent and trivalent cations.

5. The following Tables from Sigma/Aldrich product literature summarize the properties of common chelating agents. The upper table presents metal ion binding affinity in the form of $\log K_{association}$ constants. The bottom table gives the pK_a.

	EDTA	EGTA	HEDTA	NTA
Ag(I)	7.32	6.88	6.71	5.16
Ca(II)	10.96	11.00	8.14	6.41
Cd(II)	16.46	16.70	13.6	9.54
Co(II)	16.31	12.50	14.4	10.38
Cr(III)	23.40			>10
Cu(II)	18.80	17.88	17.55	12.96
Fe(II)	14.33	11.92	12.2	8.84
Fe(III)	25.1	20.5	19.8	15.87
Hg(I)	21.8	23.12	20.1	14.6
Li(I)	2.79	1.17		2.51
Mg(II)	8.69	5.21	7.0	5.46
Mn(II)	14.04	12.3	10.7	7.44
Na(I)	1.66	1.38		2.15
Ni(II)	18.62	13.55	17.0	11.54
Pb(II)	18.04	14.71	15.5	11.39
Sn(II)	18.3	23.85		
Tl(III)	22.5			18
Zn(II)	16.50	14.5	14.5	10.67

Physicochemical Data of Some Complexanes

The term complexane has been recommended by IUPAC for EDTA and other aminopolycarboxylic acids of related structure. The pK values of some complexanes are given below.
In the table above, the absolute stability constants of various metal complexes of these complexanes are reported. As mentioned above these data permit calculation the apparent stability constants of these complexes at any pH. Data were taken from reference [6].

	EDTA	EGTA	HEDTA	NTA
pk_1	1.99	2.00	2.51	1.89
pk_2	2.67	2.65	5.31	2.49
pK_3	6.16	8.85	9.86	9.73
pK_4	10.26	9.46		

EDTA	Ethylenediamine-tetraacetic acid Disodium salt
EGTA	Ethyleneglycol-O, O'-bis(2-aminoethyl)-N, N, N', N'-tetraacetic acid
HEDTA	N-(2-Hydroxyethyl)ethylenediamine-N, N', N'-triacetic acid Trisodium salt
NTA	Nitrilotriacetic acid

Concentration Units When Working with Membrane Proteins and/or Membrane Associated Ligands.

When working studying the association of molecules that are both associated with detergent micelles and/or with lipid vesicles, it is not appropriate to use bulk concentration units (molarity) to express concentrations and as units for K_d. For membrane associated molecules what matters is the concentration in the MEMBRANE.

Consider the case where you put 100 molecules of "compound A" into a single lipid vesicle that it sitting in a 1 ml solution. In this case the molecules are much more likely to bump into each other than in the case where there are 100 molecules of compound A distributed into 50 vesicles that are sitting in 1 ml of solution. While in both cases the bulk concentration is the same (100 molecules/ml), the "local" concentration in the first case (100 molecules/vesicle) is much higher than in the second case (2 molecules per vesicle).

What this means is that for meaningful thermodynamic measurements to be made regarding molecular association in micelles or lipid vesicles, you have to use concentration units that express the concentration of the molecules of interest within the membrane-mimicking phase.

Usually the units that are used are mole % units. If there are a total of 100 lipids in vesicles for every molecule of protein X then the concentration of protein X is 1/101 X 100 = 1 mol%.

$$\text{Mole \% of solute A in micelles or vesicles} = \frac{\text{number of moles of solute A}}{\text{total moles of molecules in each micelle or vesicle}} \times 100$$

For example, if you have 22 mM DPC micelles that contain 20 mol% POPC (a lipid) and you solubilize 1 mM of diacylglycerol kinase the mol% of diacylglycerol kinase would be:

molarity of DPC in micelles = 22 mM – critical micelle concentration (which is 2 mM for DPC)
= 20 mM

$$\text{mol\% DAGK} = \frac{\text{1 mM DAGK}}{\text{20 mM (micellar DPC) + 5 mM (POPC) + 1 mM (DAGK)}} \times 100 = 3.8 \text{ mol\%}$$

For molecular association between two membrane-associated molecules, K_d will therefore be expressed in mol% units.

Note that when working with micelles or mixed micelles you can assume that the mixing of component molecules between micelles will be quite rapid on most experimental time scales. However, for very hydrophobic molecules in lipid vesicles (liposomes) this is often not the case. If you have two vesicles in solution, the rate at which lipids and integral membrane proteins "hop" from one vesicle to the other may be very very slow. This must sometimes be taken into account in experimental design and data analysis.