

# MOLECULAR INTERACTIONS AND ENZYME KINETICS

© Jolyon Jesty  
(jolyon.jesty@sunysb.edu)

---

This material, although freely available, is copyright, and *neither text nor figures may be reproduced in any form without the copyright notice and attribution*. I wrote it for a 9-lecture part-course for molecular biology graduate students, which I no longer teach. However, it remains available for those who might find it useful. The material introduces quantitative analysis of biochemical interactions at a level suitable for upper-level biochemistry undergraduates, or graduate students whose training and expertise do not necessarily include biochemistry. Its emphasis is on real examples and working problems. There are four sections: 1) the analysis of molecular interactions in terms of both equilibria and rates; 2) a brief introduction to transition-state theory, and how enzymes make reactions go faster; 3) the core of enzyme kinetics, including basic kinetics and an outline of inhibitor types and mechanisms; and 4) an essential Appendix.

---

## INTRODUCTION

The basic approach and basic rules are essentially the same for the equilibrium and rate analyses of almost all molecular interactions, and they fall into two major classes:

- Solution phase: e.g.  
soluble enzyme + substrate  
soluble enzyme + inhibitor  
soluble enzyme + cofactor  
oligonucleotide + DNA-binding protein
- Mixed phases: e.g.  
membrane enzyme + soluble substrate  
soluble enzyme + membrane cofactor  
soluble ligand + membrane receptor  
chromatin DNA + soluble DNA-binding protein

While there are many reactions that occur only in a planar phase, like the interactions of membrane proteins, and structural aspects of these reactions are frequently addressed, their quantitative kinetic analysis is very hard and very rare.

In some of these examples, the binding of two or more molecules is only part of the story we can quantify. For example, the binding affinity between a soluble hormone or agonist and its membrane receptor is a major piece of quantitative information, but it hardly tells the whole story. Many hormones and other physiological effector molecules, having bound to a cell membrane, will get taken into the cell and—for instance—find their way to the nucleus and bind to specific nuclear proteins. At this point we might again become interested in measuring something quantitative like the affinity of a hormone for another specific binding protein. Then we might investigate the affinity of a molecule, such as a transcription factor, for the promoter region of a particular gene. At each stage where we might be interested in quantitative measurement of interactions, it is the *affinity* of the interaction that is usually central.

**Affinity** is another way of talking about an *equilibrium* constant, i.e. an association constant ( $K_a$ ) or a dissociation constant ( $K_d$ ). But hidden in the use of these terms, and their values, is the fact that **equilibria and affinities are defined by the balance of rates**. I will tend to separate discussion of equilibrium things from rate things, but always remember that underlying the descriptions of affinities and equilibria is a dynamic balance of rates—the rates of association of molecules, and conversely the dissociation of the combined molecules. *You cannot talk about affinity in a reaction that you don't know is reversible.*

**Enzymes** require us to go further than just talking about affinity. Here the binding of substrate(s) to an enzyme is just a precursor to the chemical reaction that the enzyme catalyzes. The first part of the process—combination with the substrate—is treated exactly the same way as we might treat any other molecular interaction; but on top of this we have to consider how

to quantify what happens to the enzyme:substrate complex...i.e. the formation of products. These are once again *rate* questions.

**Steady states** have a central role in dynamically regulating concentrations. Just for example, consider what controls the concentration of a protein in the cell cytoplasm. Many might consider this to be solely a function of protein synthesis, i.e. the cytoplasmic concentration is purely controlled by transcription and translation (expression) rates, perhaps explaining why so much effort is spent on understanding these processes. Not so. The concentration is defined by the steady state of the dynamic balance between synthesis and degradation, and as such it is *directly* controlled, and just as effectively too, by the protein's degradation rate. Steady states are also central in enzyme kinetics.

## CONCENTRATION NOTATION

I find the standard square-bracket notation for concentrations awkward to type and also, when expressions get complex, very awkward to read. I will use *italics* to denote concentrations of things, e.g. *E* means the concentration of the enzyme E. In order to avoid problems in discriminating between single molecules and complexes (e.g. between E and S separately, compared with an ES complex), multiplication of concentrations is shown, where necessary, by a raised period, e.g. where I write  $K_d = E \cdot S / ES$ , I mean, in square-bracket format,  $K_d = [E][S]/[ES]$ .

---

## MOLECULAR INTERACTIONS

We start by considering how to analyze algebraically the very basic interaction of just two molecules, either with both in solution, or with one in solution and one on a membrane. As yet they are not reacting beyond the fact of their simple affinity for each other, i.e. **they bind**.

### A REVERSIBLE BIMOLECULAR INTERACTION AT EQUILIBRIUM

#### Enzymes & Inhibitors

I will first consider two molecules, both in solution, reacting to form a single bimolecular complex. A typical example is the reaction of an enzyme, E, with a reversible inhibitor, I, but **exactly the same analysis is valid for any interaction between two molecules where at least one of them is in solution**. In particular, the situation is identical for the reversible interactions of membrane receptors and soluble ligands, as we will see later.

An **equilibrium constant is the ratio of two rate constants**, i.e. the reaction must be able to go in both directions. So an absolute requirement when talking about affinity or  $K_d$  is that the interaction being studied is reversible:



When the concentrations of the reactants get to the point that the forward and reverse (or *on* and *off*) reactions are exactly in balance, the system is **at equilibrium**. Using the subscript *eq* to denote the concentrations of the reactants at equilibrium,

$$k_{\text{on}} E_{\text{eq}} I_{\text{eq}} = k_{\text{off}} EI_{\text{eq}} \quad (1)$$

The **dissociation constant,  $K_d$** , is defined as a ratio of rate constants,

$$K_d = \frac{k_{\text{off}}}{k_{\text{on}}} = \frac{E_{\text{eq}} I_{\text{eq}}}{EI_{\text{eq}}} \quad (2)$$

A high-affinity interaction will have a large "on rate",  $k_{\text{on}}$ , and a small "off rate",  $k_{\text{off}}$ . Thus,  **$K_d$  is the inverse of affinity**. Unlike biochemists, chemists usually use the *association* constant,  $K_a$ , when they talk about equilibrium constants:  $K_a = 1/K_d$ .

This difference between disciplines will crop up again when we look at the mechanisms of enzyme catalysis and  $\Delta G$ ...it's important to remember it.

For bimolecular interactions of *macromolecules* in biochemistry, a  $K_d$  of 1  $\mu\text{M}$  ( $10^{-6}$  M) is a fairly low affinity, 1 nM ( $10^{-9}$  M) is getting pretty potent, and 1 pM ( $10^{-12}$  M) is a very high affinity. Anything beyond that ( $K_d < 1$  pM) is essentially irreversible. **In general, variations in affinity are caused more by differences in  $k_{\text{off}}$  than  $k_{\text{on}}$ .** Over the range mentioned, dissociation (off) rates may vary from half-lives of a few seconds at  $K_d = 1$   $\mu\text{M}$  to days at  $K_d = 1$  pM.

### Measuring $K_d$

Let us start with a mixture containing an enzyme and inhibitor, E and I. You must first remember that the concentrations of reactants in Eq. 1 and 2 above are concentrations at equilibrium. If you start with known amounts of E and I and let them react, their concentrations will change as the EI complex is formed. The analysis that follows allows us to **express the equilibrium concentrations of E and EI if we know the starting concentrations of E and I**, called  $E_0$  and  $I_0$ , and the value of the  $K_d$ . For enzyme inhibitors,  $K_d$  is generally called the inhibitor constant, or  $K_i$ , so we will use that term in this example.

### Analysis: the algebra

**To measure  $K_i$  experimentally** we will take a fixed concentration of E—call it  $E_0$ —and do a series of incubations at increasing concentrations of I. In each incubation we allow the reaction to come to equilibrium, and then measure the concentration of free (uninhibited) E, called  $E_{\text{eq}}$ . What we then need for analysis of the data is an equation that expresses this **measured  $E_{\text{eq}}$  as a function of the varying concentration of I** and the value of the  $K_i$ . For the moment we will assume that we either know, or can measure, the concentration of I at equilibrium,  $I_{\text{eq}}$  (this will be discussed further below). We start by rearranging Eq. 2 (and using  $K_i$  for this example instead of  $K_d$ ) to get

$$E_{\text{eq}} I_{\text{eq}} = EI_{\text{eq}} K_i \quad (3)$$

The one thing we don't know or measure here is  $EI_{\text{eq}}$ , which has to be got rid of. Since E can be present as either free E or EI,

$$EI_{\text{eq}} = E_0 - E_{\text{eq}} \quad (4)$$

Substituting this in Eq. 3, we get

$$E_{\text{eq}} I_{\text{eq}} = EI_{\text{eq}} K_i = E_0 K_i - E_{\text{eq}} K_i,$$

which, rearranged, gives

$$E_{\text{eq}} = \frac{E_0 K_i}{K_i + I_{\text{eq}}} \quad (5)$$

If  $EI_{\text{eq}}$  is required rather than  $E_{\text{eq}}$ , we see that  $E_{\text{eq}} = E_0 - EI_{\text{eq}}$  (Eq. 4), and substitute that into Eq. 3 instead, to get

$$EI_{\text{eq}} = \frac{E_0 I_{\text{eq}}}{K_i + I_{\text{eq}}} \quad (6)$$

**Eq. 5 and 6, and their equivalents for any other bimolecular equilibrium systems, are fundamental in the study of just about any reversible interaction.** However, they are only of use if we know or can measure  $I_{\text{eq}}$  (or its equivalent in other systems). For people studying enzymes,  $I_{\text{eq}}$  cannot usually be measured, and experiments are therefore set up so that it is known by assumption rather than measurement. The commonest, and usually easiest, way to do this is to **set  $I \gg E$** , so that the formation of the EI complex will not significantly affect the concentration of I. Thus

$$I_{\text{eq}} \sim I_0 \quad (7)$$

### Receptors & Ligands

Studying ligand-receptor interactions,  $R+L \rightleftharpoons RL$ , is rather easier. Let us rewrite Eq. 6 in terms of R and L:

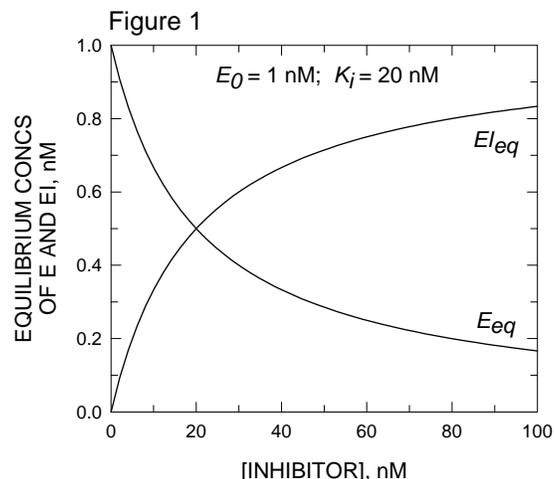
$$RL = \frac{R_0 L_{eq}}{K_d + L_{eq}} \quad (8)$$

In this case the experimenter, who is usually using a labeled preparation of the ligand, can measure  $L$  at equilibrium both in solution ( $L_{eq}$ ) and on the cell ( $RL_{eq}$ ). This is a particular advantage, since she no longer has to assume that  $L$  is in large excess over  $R$ ; in fact, as long as  $L \geq R$ , any experimental setup is feasible.

## Enzymes & Inhibitors again

Having shown how easy—at least in this regard!—the analysis of receptor experiments might be, let us return to enzymes and inhibitors to look at the forms of Eq. 5 and 6.

**Eq. 5 describes the reduction in free E at equilibrium as  $I$  increases, while Eq. 6 describes the corresponding increase in EI.** Examples are shown in Fig. 1, for an arbitrary enzyme, at 1 nM, being inhibited by an inhibitor with a  $K_i$  of 20 nM. At infinitely high  $I$ ,  $E$  eventually goes to zero because all E is present as the EI complex. Note, however, that the levels of  $I$  required to drive the equilibrium to near-total inhibition are very high. The reduction in E depends on the value of the term  $K_i/(K_i + I_0)$ . For instance, at 100 nM ( $5 \times K_i$ ) this has a value of 1/6, so 17% of the uninhibited enzyme is still present.



You may also note that in this extremely simple system, containing only E and I, 50% enzyme inhibition is observed when  $I = K_i$ . **This is true only if no other reaction is interfering.** If, for example, a substrate of E were present, the  $E+I \rightleftharpoons EI$  equilibrium would be affected, and would change the level of  $I$  required for 50% inhibition. I will discuss this in more detail under Enzyme Inhibition.

## Units

Although Fig. 1 shows both enzyme and inhibitor in molar terms, the analysis so far does not actually require enzyme molarities (or receptor numbers in receptor-ligand systems) to be known. Consider the following examples:

**$E+I \rightleftharpoons EI$**  (Eq. 5 or 6). Here,  $E$  and  $EI$  must be in the same units, but  $I$  can be in different units. The results of an experiment would be estimates of (1)  $E_0$ , in the concentration units of  $E$ ; and (2)  $K_i$ , in the concentration units of  $I$ . This is common when crude preparations of enzyme of unknown molar concentration are used for enzyme studies.

**$R+L \rightleftharpoons RL$**  (Eq. 8) In this case  $R$  and  $RL$  are in the same units (e.g. number of receptors or occupied receptors), but  $L$  can be different (e.g. the solution concentration of  $L$ ). This does not happen often in practice, because  $RL$  and  $L$  will be measured in the same units anyway.

## Experimental considerations, concentration ranges

**In order to measure a  $K_i$  or  $K_d$ ,** the inhibitor or ligand concentrations ( $I_{eq}$  or  $L_{eq}$ ) should usually range from about  $K_i/4$  to  $4 \times K_i$ , with equal numbers of points above and below the value of  $K$  (i.e. not equally spaced over the range). Also, remember that in the case of enzyme inhibitors, where  $I_{eq}$  cannot be directly measured, analysis using Eq. 5 or 6 requires that  $I_0$  substantially exceed  $E_0$  over the whole experimental range.

## High-Affinity Interactions, E+I

From this rough rule, the experimenter in Fig. 1 needed to measure  $E_{eq}$  over the approximate range  $I = 5 - 80$  nM. The concentrations of  $E_{eq}$  that she would have been measuring were in the range roughly 0.15 - 1 nM. Her problem is that, while a  $K_i$  of 20 nM is nothing unusual, *many enzyme assays are insufficiently sensitive* in the subnanomolar range. But she cannot raise  $E_0$  much higher, because then she is no longer meeting the requirement that  $I_0 \gg E_0$ . What to do?

The root of this problem is having to set up the experiment so that  $I_{eq} \sim I_0$  (Eq. 7) This is what enables the use of Eq. 5 or 6. What happens if we don't set  $I \gg E$ ? We can account for the reduction in  $I$  when EI is formed, by including the following in the derivation:

$$I_{eq} = I_0 - EI_{eq} \quad (9)$$

When this complication is included, deriving an expression for  $E_{eq}$  requires the solution of a quadratic:

$$E_{eq} = \frac{1}{2} \left( E_0 - I_0 - K_i + \sqrt{(E_0 - I_0 - K_i)^2 + 4K_i E_0} \right). \quad (10)$$

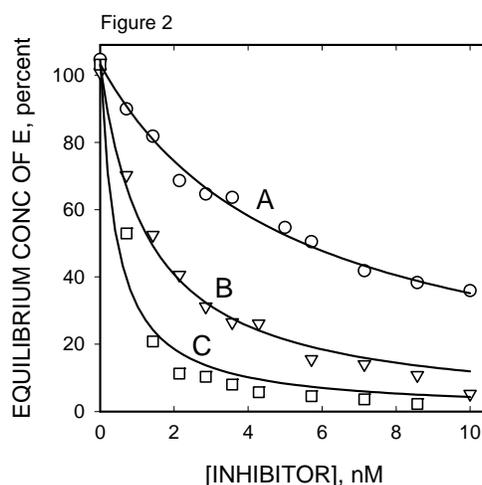
Once again,  $E_{eq}$  can be obtained if it is the EI complex that you happen to be measuring.

These days data can be easily fitted to functions like this by nonlinear regression (see *Appendix*). It gives a substantial experimental advantage. Now you may wonder whether it is ever necessary to use the simpler functions, Eq. 5 and 6. Why not always use Eq. 10? Well, for a start it is not necessary for receptor-ligand studies, where  $L_{eq}$  is usually easily measured, and no assumptions like Eq. 7 are necessary about it. Second, in enzyme and similar studies, it has a significant disadvantage, that everything [i.e.  $E_{eq}$  (measured), and  $E_0$  and  $I_0$  (both fixed by you)] **must now be in the same molar units**. In practice this means you usually need enzyme of known molar concentration in your assay, i.e. the assay must at some point have been standardized with a pure enzyme.

## Recognizing High-Affinity Problems, E+I

As working biochemists deal with interactions of ever higher affinities, a major problem can arise: that it is much too easy to analyze an experiment using the standard simple functions (Eq. 5 or 6) and not realize that they are the wrong functions. Fig. 2 shows an example, of the inhibition of an enzyme E by **three reversible inhibitors with increasing affinity (A, B, C)**. Our experimenter here is checking three new inhibitors. Because the enzyme is not yet stable when it is purified, and the protein folks have not yet fixed that problem, she is having to use a crude preparation of only roughly known molar concentration.

[Kinetic studies of crude enzyme preparations are common... this is not an arbitrary teaching fiction. As long as the main assumption of Eq. 5 or 6 is met ( $I \gg E_0$ ), the unit used for  $E$  and  $EI$  doesn't matter (see above). It is therefore fully valid, when using Eq. 5, to express  $E$  as a percentage of maximum, as here.]



The inhibition data for the three inhibitors are shown in Fig. 2, along with the best-fit lines our colleague obtained when she fitted each data set to Eq. 5. The best-estimate  $K_i$  values obtained from the fits shown (see *Appendix* about fitting) were A) 5.2 nM; B) 1.3 nM; C) 0.43 nM. Just by looking at the way the points distribute about the best-fit lines, it is clear that there is a problem for C, and probably for B too—in both cases, early points are above the best-fit line, and late points below it. **For B and C the function (Eq. 5) does not describe—is wrong for—these data.**

**The real conditions of Fig. 2.** These data actually represent the results of experiments containing **1 nM enzyme**. Because of the requirement that  $I > E_0$ , any data below about 2 nM inhibitor would be suspect, and that indeed is the range where problems are seen. The approximate  $K_i$  values here are in fact: **A) 5 nM; B) 1 nM; and C) 0.2 nM**. The number for **A** is determined well because there is plenty of usable data up to 10 nM inhibitor. Errors in the early data will be minor because there is little inhibition (and therefore little reduction in  $I$ ) from 0 to 2 nM, and they won't have any big effect on the estimation of  $K_i$ . For **B**, 50% inhibition is reached at just over 1 nM inhibitor, so the important data in the first part of the curve, which are in error, significantly affect the estimate of  $K_i$  (30% error). In curve **C**, 90% inhibition is reached at 2 nM inhibitor, so almost *all* the important data (from 100% to 20% activity) is in error because  **$I_{eq}$  in this range does not approximate  $I_0$** . And unfortunately, because she is measuring enzyme activity in arbitrary (not molar) units, our experimenter cannot use Eq. 10. The result is that the  $K_i$  for C is overestimated by more than 2-fold, i.e. the inhibitor is more than twice as potent as the analysis suggests. I think she is going to have to purify the enzyme, or at least determine its molar concentration in the crude preparation....but at least she noticed the problem!

## Very High Affinity: From Equilibrium to Stoichiometry

Let me continue this story to its conclusion, of what happens when you have a **very-high-affinity interaction**, and how methods devised for equilibrium studies become useless...sometimes again without your realizing it. Fig. 3 continues the same story as Fig. 2, starting with 1 nM enzyme, but the affinities are getting higher and higher. At a  $K_i$  of 1 nM, there is plenty of data there about the  $E+I \rightleftharpoons EI$  equilibrium, and the  $K_i$  will be reasonably well determined. At the **lower  $K_i$  values**,

large parts of the plots are straight; in fact at  $K_i = 0.01$  nM, the line is dead straight down to about  $E_{eq} = 0.2$ . **This is a titration line.** In other words, the observations are about **stoichiometry**, not equilibrium affinity. Since there is no equilibrium information in a straight line, affinity cannot even be guessed. For example, 50% inhibition for the two lines to the left is attained at  $I = 0.5 - 0.6$  nM. This has absolutely nothing to do with affinity or  $K_i$ ; it simply shows (in this example) that the affinity is very high, and that **the enzyme:inhibitor stoichiometry is 1:1.**

### Beware log(dose) transforms

In Fig. 3 the straight-line—i.e. non-hyperbolic—nature of the two high-affinity plots is clear, and they would quickly alert our experimenter that these data could not be used to determine  $K_i$ . However, other plots are commonly used, in particular the use of log-transformed X (dose) axes, as advocated by **Klotz** and others. Really not recommended...

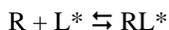
Fig. 4 shows the **identical data to Fig. 3** plotted on a log-scaled X axis (except the points at  $I = 0$  are missing: they cannot be plotted on a log axis). The important linear, or titration, nature of the two high-affinity lines has disappeared, and all three look more or less sigmoid. It is tempting to quickly read off the inhibitor concentrations that produced 50% inhibition, and assume that these represent the affinities. In the case of the left two lines, this is hopelessly incorrect (see the actual numbers on the plot). This type of plot is commonly found in all sorts of areas in biology and biochemistry. By themselves they are much too prone to misinterpretation when affinities are high (and don't forget you might well not know that).

### How do you check that your conclusions about affinities are valid?

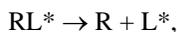
**1) First**, you fit, for instance, to Eq. 5; plot your data as plain hyperbolas; and compare the fitted line to the data points. The two lefthand data sets in Fig. 3 clearly do not fit the simple hyperbola, and it will show. **2) To prove** whether or not there is a problem, repeat the experiment at a different enzyme concentration. An intrinsic property like affinity— $K_i$ , for example—does *not* change with enzyme concentration; but stoichiometric results change in proportion. This is the key test.

### Reversibility, e.g. Ligand-Receptor Interaction

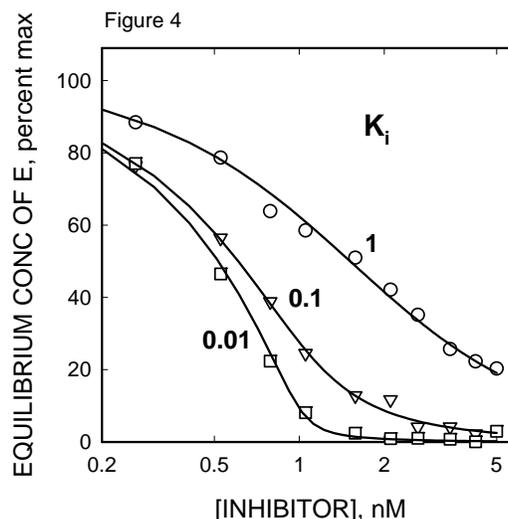
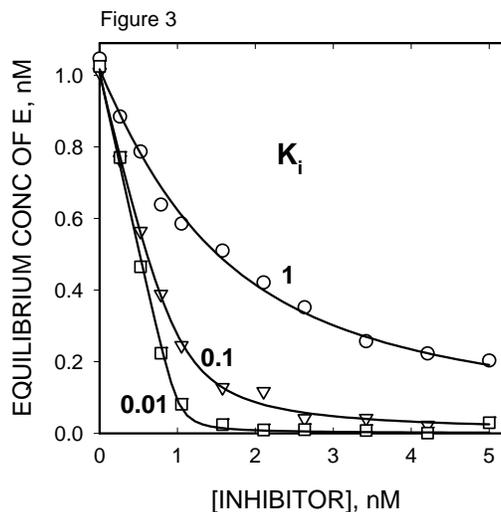
Let us return to using receptors and ligands as our example of a reversible bimolecular reaction. First, a reminder that whenever you use the word **affinity** you are talking about a **reversible** interaction. Although many do not bother, it is usually good—particularly when binding ligands to receptors, but also in other areas—to check that your reaction actually *is* reversible, i.e., that you can make it go backwards, and perhaps measure the rate of that reaction,  $k_{off}$ . Here is the standard method. First you let cells (receptors) and ligand react to attain equilibrium, just as you did to measure the  $K_d$ :



where  $L^*$  represents radiolabeled (or other suitably labeled) ligand. You take a sample at this point and measure  $RL^*$  (i.e. bound  $L^*$ ) and free  $L^*$  in the supernatant. You then add a large excess of unlabeled ligand,  $L$ . Consider now what happens when  $RL^*$  dissociates. First,



freeing up receptors. However, any available (unoccupied)  $R$  will now pick up *unlabeled* ligand rather than *labeled*, because the unlabeled is now in very large excess (typically >100-fold), i.e.



The result is that as  $L^*$  dissociates from R, it never returns. So **the rate of dissociation of labeled  $L^*$  is a direct measure of  $k_{\text{off}}$** . Experimentally, timed samples are taken, and the amount of  $RL^*$  remaining on the cell is determined in each. The kinetics are extremely simple ( $k_{\text{off}}$  is a first-order rate constant), and described by

$$RL_t^* = RL_0^* \cdot e^{-k_{\text{off}}t} . \quad (11)$$

Eventually all  $L^*$  should dissociate, but in practice—particularly when working with cells, which often internalize receptor-bound ligands—this requirement is often not met. If you know for certain that some of your "bound" ligand is in fact internalized  $L^*$  that is nondissociable (call it  $RL_{\text{nd}}$ ), you may fit data to a single exponential that has a nonzero asymptote:

$$RL_t^* = RL_{\text{nd}}^* + RL_{\text{diss}}^* \cdot e^{-k_{\text{off}}t} . \quad (12)$$

## Microplate Solid-Phase Binding Assays (and Abioco)

### Reversibility: show it

An exactly equivalent setup to  $R+L \rightleftharpoons RL$  is the measurement of binding affinities between two proteins when one of them is bound to a microplate (another name is ELISA plate). One component is bound to the wells of a plate, and this is physically—and in analysis—equivalent to the membrane-bound receptor on a cell, R. The other component is added in solution to the plate, allowed to come to equilibrium, and the amount bound to the plate (equivalent to RL) is then measured. This method has all the possible problems that I have already described (e.g. Fig. 2-4), and then some more. Investigators only rarely check the **reversibility of binding**, though it is easy enough to do. But the main problem is the assumption that the protein bound to the plate has the same properties as it does in solution. It is quite feasible that **binding a protein to reactive polystyrene** affects its interaction with the solution protein,  $k_{\text{on}}$ . Moreover, once the protein:protein complex is formed on the plate, we also do not know whether its dissociation rate,  $k_{\text{off}}$ , is equivalent to what occurs in solution. Remember *it is the ratio of these two rates that defines the  $K_d$* . A couple of careful studies of methods in the literature suggest that microplate affinity assays can be subject to huge errors, sometimes of *several orders of magnitude*. Given these problems, they are probably best used only as confirmatory methods.

Note added in the 21st century... There is a large, very expensive machine, called the Abioco, which is essentially a means of following the kinetics of binding of ligands to surface-bound proteins. However, the "surface" here is exactly equivalent to a microplate: a reactive piece of surface that is treated to bind proteins. All that is gained is *rate* information, which you cannot get from microplates. The results generate just the same concerns as the more mundane microplate reader. The only Abioco information I have any trust in is studies that compare and report studies that look at a reaction both ways: [A (chip-bound) + B (solution)] **and** [B (chip-bound) with A (solution)]. If they don't agree, forget it...

### Washing off excess ligand: losing the bound?

A closely related problem is that ELISA binding methods require the near-complete removal of the solution of ligand from the wells of the plate before measuring the amount of bound ligand. (This is not true of the Abioco, which uses very fancy optical methods to measure bound protein on the solid surface.)

[In cell-receptor work much more efficient separation of cells from solution is the norm: one of the best ways is to centrifuge the cells through an oil that is less dense than the cells and more dense than the solution. The cell pellet is then just snipped off and counted. Carryover of solution ligand with the cells in this case is minuscule. Rapid filtration with a fast wash of 1-2 seconds is also used: the cells, with bound ligand, are trapped on the filter.]

To remove excess ligand solution from a microplate requires washing, and plate washing is slow (many seconds, minimum). If the affinity is weak, it is quite feasible that over the time you are washing the plate, the ligand is dissociating from the protein bound to the well while you are doing the wash. This should always be checked, and is easy to do: simply compare the bound counts obtained from a quick wash with those from a slow wash. If they are different, you have a problem; if not, be happy.

## MULTIPLE (PARALLEL) INTERACTIONS

To this point I have focused entirely on the interactions of two molecules. More complex, and more common, problems involve multiple, parallel, interactions. I will stick with the ligand-receptor example for a while, because that is a common place to examine the reaction, for example, of a single ligand with two cellular receptors.

### Nonsaturable sites

The most common situation is that a cell, in addition to having a "specific" receptor for a ligand, also binds it at sites that are

present in much larger numbers, but with **much lower affinity**. These are often called "**nonspecific**" interactions, although it is quite feasible that they might be just as specific as the interaction that you are interested in. "Nonspecific" is a subjective word usually used to mean an interaction that *you* are not interested in. A better term is **nonsaturable**: the result of the lower affinity is that binding continues to increase with ligand concentration even up to high levels, i.e. the binding does not saturate. In other words, binding to such sites is simply proportional to the solution ligand concentration,  $RL = C \cdot L$ , where  $C$  is a proportionality constant. If a cell has a saturable class and a nonsaturable class of receptors, **the total ligand bound (and measured) is simply the sum**:

$$(Bound L)_{eq} = \frac{R \cdot L_{eq}}{K_d + L_{eq}} + C \cdot L_{eq} , \quad (13)$$

where  $R$  is the saturable receptor. If you are only interested in  $L$  binding to  $R$ , what is the best way to handle nonsaturable interference? The poor approach is to do a measurement or two at much higher concentrations of  $L$ , *assume that you have totally saturated your high-affinity receptor*, and simply subtract the "nonspecific" number. Based on a few points and an assumption, this in effect gets you a value for  $C$ . A rather better approach is indeed to extend your series with a few observations at considerably higher concentrations of  $L$  (e.g.  $L \sim 10, 20, \text{ and } 50 \times K_d$ ), but then fit *all* the data directly to Eq. 13.

Fig. 5 shows three sets of data for a ligand binding to a single saturable site, plus three levels of nonsaturable interaction, shown by increasing values of  $C$ . As you might guess,  $R_0$  and  $K_d$  are usually determined less accurately as  $C$  becomes large (note, by the way, that this shows a very poorly designed experiment, with insufficient data below the  $K_d$ ...hence the large errors. It's not so easy making convincing fake data!). For the data shown, fitting to Eq. 13 gave:

$C = 0$ :	$R_0 = 2.25 \pm 0.63$ ;	$K_d = 12.8 \pm 6.6$
$C = 0.02$ :	$R_0 = 2.02 \pm 0.60$ ;	$K_d = 10.5 \pm 6.3$
$C = 0.2$ :	$R_0 = 2.40 \pm 0.51$ ;	$K_d = 15.1 \pm 5.4$

In reality (plus some added noise), the numbers were  $R_0 = 2$ , and  $K_d = 10$ .

## Multiple saturable sites

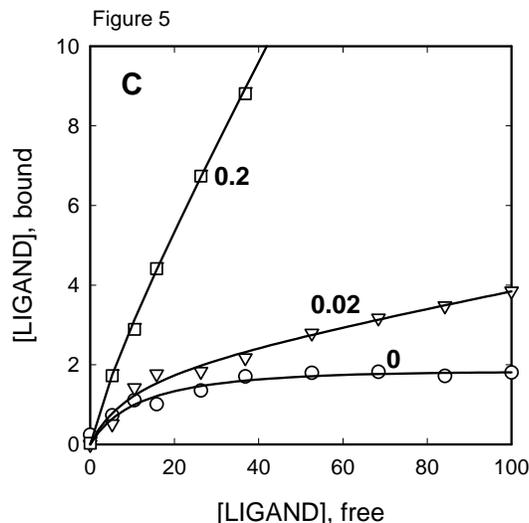
Sometimes you will suspect, or have other initial evidence for, more than one saturable interaction of your ligand with a cell, i.e. more than one class of receptor. (Or, in enzyme territory, you may suspect two isoforms of an enzyme, which have different affinities for a substrate or inhibitor.) There are many papers in the literature where *proof* of this is largely unsupported by the data. With multiple receptor sites, **the total ligand bound is again the sum**. For two sites,

$$(Bound L)_{eq} = \frac{R_1 L_{eq}}{K_1 + L_{eq}} + \frac{R_2 L_{eq}}{K_2 + L_{eq}} , \quad (14)$$

where  $R_1$  and  $R_2$  are the receptor numbers (concentrations) for the two sites, and  $K_1$  and  $K_2$  are the corresponding dissociation constants. If (as is common) nonsaturable interactions are seen, you add a  $C \cdot L_{eq}$  term to this as in Eq. 13. That makes **five unknowns** that you are going to be solving for when you fit the data: **you need huge and accurate data sets for this stuff**.

If the  $K_d$  values are close, then obviously the observed binding will appear experimentally as one site with a single  $K_d$ , the number of receptors simply being the sum of the two. **In order to prove the existence of two independent sites, the  $K_d$  values must, even with near-perfect data, differ by at least 5-fold**. Even with a 10-fold difference in  $K_d$ , a clear conclusion will often be difficult. The **Appendix** includes a description of a set of data derived from a 2-site model that upon analysis didn't support the existence of two sites.

As you might guess, ever-more-complex expressions can easily be written for yet more sites. Anyone who even pretends to fit experimental data to such things needs their head looking at...read about **Occam's Razor** in the Appendix.



## Allosteric Interactions

This area is too complex for a proper description here. The term *allosteric* means a change in the kinetic characteristics or ligand-binding characteristics at one site on a protein caused by the binding of a ligand molecule at another site. The different sites may bind different molecules (e.g. a substrate at one site on an enzyme; an allosteric regulator at another), or the same molecules (e.g. substrate binding to an enzyme, or oxygen to hemoglobin). In the latter case, where a protein binds more than one ligand molecule, the behavior leads to cooperativity, i.e. the binding of each ligand or substrate molecule changes the binding for subsequent ones. Affinity may increase (in which case the word is cooperativity) or decrease (for which the phrase is negative cooperativity). A central feature of allostery is that **it cannot be explained by the existence of multiple independent sites** of different affinity: if that were so, the high-affinity sites would simply get occupied first, followed by the lower-affinity. It depends on **interaction between sites**, so that site occupancy is the controlling factor. "Allostery" really doesn't include cooperative behavior between different molecules, but it is worth noting that the kinetic behavior is very similar, since the properties at a vacant site are changed by the occupancy of other site(s).

Cooperative effects include:

- the binding of ligands to cellular receptors, where occupancy of one receptor site may reduce the affinity for the ligand at unoccupied sites. This is negative cooperativity, and it is a common means of "down regulation" of receptors.
- positive cooperativity in ligand-protein interactions is found in molecules like hemoglobin, where occupancy of early sites greatly increases the affinity for binding to later sites, i.e. Hb(O<sub>2</sub>)<sub>3</sub> has a much higher affinity for O<sub>2</sub>—more than two orders of magnitude higher—than deoxygenated Hb does.
- binding of regulatory molecules to key enzymes in the positive or negative control of metabolic pathways by upstream precursors or downstream products.

## The Hill Plot: Analysis of Cooperativity

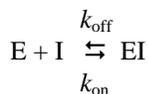
The **Hill Plot** is an empirical measure of cooperative interactions. It does not address possible molecular mechanisms of cooperativity. The system being studied may be a ligand binding to a cellular receptor, or a ligand binding to a soluble molecule (e.g. O<sub>2</sub> to hemoglobin), or a substrate to an enzyme: in the latter case the measurement will be a kinetic one rather than a direct measurement of binding. For a ligand-binding experiment, using  $L$  and  $L_{\max}$  to denote the amount of ligand bound and the amount bound at saturation,  $\log(\text{occupied/unoccupied})$  is plotted against  $\log(L)$ :

$$\log\left(\frac{L}{L_{\max} - L}\right) = h \log(L) - \log(K),$$

where  $h$  is the **Hill constant**, which is a measure of the degree of cooperative behavior, and  $K$  is a dissociation constant ( $K_d$ ). Analogous equations are used for other measures of the saturation ratio, e.g. for a cooperative enzyme the plot would be  $\log[v/(V_{\max} - v)]$  against  $\log(S)$ . If the system shows cooperativity, the plot is sigmoid with a maximal slope at about 50% saturation. In this middle region (say, between 20-80% saturation) the line is approximately straight and the slope is  $h$ . For positive cooperativity (affinity increases with increasing amount of ligand bound),  $h > 1$ , its value depending on the extent of cooperativity. For negative cooperativity (affinity decreases),  $h < 1$ . For noncooperative systems,  $h = 1$ .

## RATE STUDIES

We will not revisit equilibria until we get to enzymes and catalysis. Remember that **equilibria are the result of a dynamic balance between opposing rates**. Sometimes these rates are extremely rapid, as we often see in the interaction of enzymes with substrates and many inhibitors. Others are not. Binding of ligands to cells, for instance, often proceeds quite leisurely with, for instance, 10 or 15 minutes being required for equilibrium to be attained. Similarly, sometimes, for antigens binding to antibodies. Another area where rates are important is inhibitors that bind very tightly, but are studied or used at very low concentration, which reduces the rate. For instance, the potency of physiological (protein) protease inhibitors is generally stated in terms of their rate of inhibition of their target enzymes at a given concentration. Let us return to enzymes and inhibitors for a discussion about rates and how we measure them.



Imagine the forward rate constant for a particular reaction,  $k_{on}$ , is fairly fast (but not excessively so), at  $10^7 \text{ M}^{-1}\text{s}^{-1}$ . To measure it, we need to measure how fast enzyme activity disappears at known inhibitor concentration(s). The total **rate equation** for E is the sum of its removal (negative) and re-formation (positive):

$$dE/dt = -k_{on}I \cdot E + k_{off}EI \quad (15)$$

If it is feasible it is best to set up experiments so that  $I \gg E_0$ , so that the concentration of I does not change significantly as the reaction proceeds. The conditions are then **pseudo-first-order**, and

$$dE/dt = -k \cdot E + k_{off}EI, \quad (16)$$

where  $k = k_{on}I$  and is a first-order rate constant.

### Decay to Zero (Irreversible)

If a reaction is irreversible or nearly so ( $k \gg k_{off}$ ), then the  $k_{off}$  term can be ignored, and we get

$$dE/dt = -k \cdot E.$$

Integration gives us the well known single exponential of a first-order decay:

$$E_t = E_0 e^{-kt}. \quad (17)$$

If  $EI$  rather than  $E$  is being measured, the single exponential instead rises to an asymptote

$$EI_t = EI_{max}(1 - e^{-kt}), \quad (18)$$

where  $EI_{max}$  equals the amount of initial enzyme,  $E_0$ . (The *Appendix* covers the properties of simple exponentials.)

But what conditions do we need in order to actually be able **to measure the reaction experimentally**? A  $k_{on}$  of  $10^7 \text{ M}^{-1}\text{s}^{-1}$  means that at  $10^{-7} \text{ M}$  inhibitor, enzyme will be inhibited at a rate of  $1 \text{ s}^{-1}$ , i.e. a half-life of about 0.7 s. (See the *Appendix* about half-lives.) This is too fast for an ordinary manual experiment on the lab bench, though easily in the range of what a rapid-reaction instrument can handle. If you want to do it with manual sampling and assays, you have to slow it down, by reducing the inhibitor concentration, yet still holding  $E_0 \ll I$ . At an inhibitor concentration of 1 nM ( $10^{-9} \text{ M}$ ), the half-life will be about 70 seconds, and if you have a rapid assay that is perhaps manageable. But is your assay sensitive enough to measure enzyme at this level? Remember that  $E_0$  must be less than 1/10 of this, i.e. 0.2 nM. To follow its decay accurately, the assay must be good to about 1/10 of this. So now we need an assay that is good down to 20 pM. That is not so common.

### Physical Detection Methods

This is a good place to mention other detection methods; I have been belaboring enzyme assays for a while. Many physical detection methods are available to follow the formation of macromolecular complexes, particularly **fluorescence**. A natural fluorophore might be used (e.g. tryptophan in a protein), or a fluorescent tag may be attached to the molecule. Changes in fluorescence emission frequently accompany the formation of a macromolecular complex because of changes in the fluorophore environment and therefore changes in its quenching. Changes can be either increases or decreases in fluorescence emission, and the time course of the change may be analyzed by fitting to appropriate exponential functions (e.g. in the simplest case, Eq. 17 or 18). Physical detection methods are also ideal for rapid-reaction studies, because no sampling is necessary as it usually is in enzyme assays or in the measurement of bound ligands.

### Decay to Equilibrium

More common than irreversible decay is a **reversible reaction**, which will eventually go to equilibrium rather than to zero, i.e.  $k_{off} > 0$  in Eq. 15. To describe the time course of disappearance of E (or its equivalent in other equilibrium systems), we integrate Eq. 16. The resulting function, though not something you would want to memorize, is useful in all sorts of practical situations (remember that  $k$  is a pseudo-first-order rate constant, e.g.  $k_{on}I$ ):

$$E_t = \frac{E_0 \left( k \cdot e^{-(k+k_{off})t} + k_{off} \right)}{k + k_{off}} \quad (19)$$

Fig. 6 shows such decays at **varying on rates, and a fixed off rate** ( $k_{\text{off}} = 0.01 \text{ s}^{-1}$ ). Let us confirm some intuitive things first:

1) When  $k = k_{\text{off}}$ , equilibrium should be attained at 50%, when the rates balance. That is true: when  $k = k_{\text{off}} = 0.01 \text{ s}^{-1}$ , the final equilibrium level of  $E$  is 0.5. Correct.

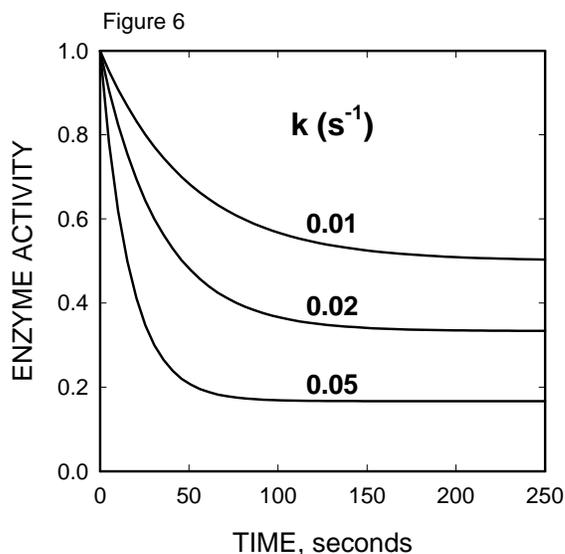
2) As the forward rate  $k$  increases (with  $k_{\text{off}}$  fixed), the equilibrium is further to the right, i.e. less  $E$  is finally present. Correct.

3) The rate at which equilibrium is attained also should go up as  $k$  rises. Correct: equilibrium is attained more quickly at higher  $k$  values.

It is also useful to see what we can get from just **examining the equation**:

1) If we set  $t = \infty$ , the exponential term ( $e^{(-)}$ ) goes to zero, and we see that the final equilibrium concentration of  $E$ ,  $E_{\infty} = E_0 \cdot k_{\text{off}} / (k + k_{\text{off}})$ ...correct from our knowledge of equilibria. And if  $k_{\text{off}} = 0$ , we have no  $E_{\infty}$  at all...correct.

2) The rate of attainment of equilibrium is controlled by the exponential ( $e^{(-)}$ ) term. This is obviously true for increasing values of  $k$ . What is not shown in Fig. 6, and it's not really intuitively obvious either, is that  $k_{\text{off}}$  also contributes: *it is the sum of the forward and reverse rates,  $k + k_{\text{off}}$ , that defines how fast the reaction gets to equilibrium.*



## If one reactant is not in large excess

If one reactant (it might be  $I$  binding to  $E$ ,  $L$  binding to  $R$ , or just protein to protein) is not in large excess, things are more complex. There are functions that describe the system in this case, and they will produce estimates of both the initial reactant concentrations and the rate constant, but they are far from perfect. For accurate work one is basically stuck with having to set things up to get pseudo-first order kinetics, i.e. one reactant is in large excess.

## More Complex Cases

Rate studies that involve more complex cases, such as more than one exponential, are discussed briefly when I consider "irreversible" enzyme inhibitors.

## Diffusion-Limited Rates

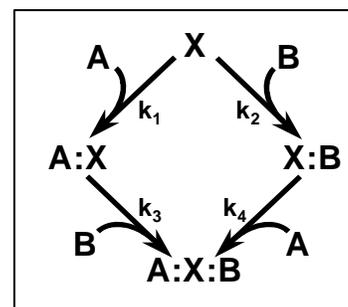
This is a techno-phrase one hears quite often, and sometimes incorrectly. In the reaction of two molecules, there is a **maximum collision rate** that depends on the statistics of collision, and therefore on **concentrations and diffusion rates**. (The foundation here is due to **Boltzmann**...a life that all scientists should read about and ponder.) If every collision results in a reaction, the rate is said to be diffusion-limited. Such reactions occur only when the geometry of collision is irrelevant, i.e. it doesn't much matter which piece of the molecules is actually involved in the collision. An example is the reaction of  $\text{H}^+$  and  $\text{OH}^-$  to form  $\text{H}_2\text{O}$ , which occurs at a rate on the order of  $10^{11} \text{ M}^{-1}\text{s}^{-1}$ . Macromolecular reactions (2 large molecules; or 1 large + 1 small) never come anywhere close to this number.

It is common to claim that bimolecular reactions of macromolecules occur at "diffusion-limited rates" of the order of  $10^8$ - $10^9 \text{ M}^{-1}\text{s}^{-1}$ . Although this is indeed about as high a rate constant as you will ever see for a reaction involving a macromolecule, it is not limited only by diffusion; rather, a very large proportion of collisions are unproductive (binding does not occur upon collision) because of **geometry, or steric, requirements**. If a particular site or orientation of a macromolecule is involved in binding to another molecule (e.g. a substrate binding to an enzyme, or a ligand to a receptor), then the great majority of collisions with the other reactant will be unproductive.

## Detection of Pathway Intermediates...A Brief Tangent...

A common problem in many different areas is the detection of intermediates on reaction pathways. An important subset of this type of problem that will serve as an example occurs when there are alternative pathways to a product, and you are trying to determine which pathway is the relevant, or more efficient, one. Imagine that you have a molecule  $X$  (e.g. a protein, or DNA, or whatever) that undergoes two modifications in either order, as here. How do you find out which is the preferred

pathway to A:X:B? Imagine you do an experiment and find that *intermediate A:X is seen at detectable levels as A:X:B is being formed, but X:B is not detected*. **This tells you nothing about the preferred route.** You then decide to separate the A and B routes by measuring the rate of disappearance of X in the presence of A only, and then with B only. In this particular experiment you find that the rates are equal, telling you that 50% of X must go down each pathway. So why does A:X build up, but X:B does not? The observation tells you that the second step of the left pathway is slow, i.e.  $k_3 \ll k_4$ . In this imagined example, then, it is the righthand pathway—with no intermediate detected—that is the more efficient: material being processed by the lefthand pathway will get there in the end, but more slowly. The lesson is that the detection of intermediates on a given pathway is not *any* indication that it is the preferred pathway.



## CATALYSIS

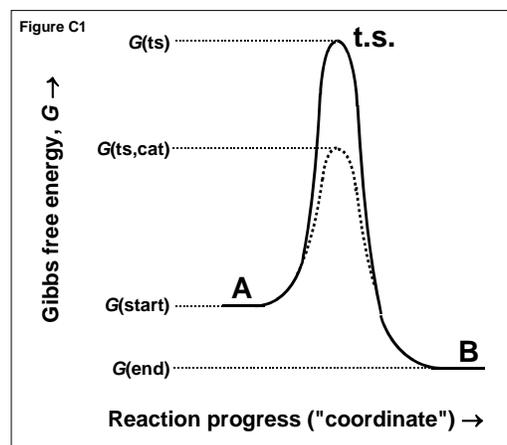
Finally we get to enzyme kinetics; but before addressing the analysis of enzyme action, we will ask how enzymes make reactions go faster. This is basically a thermodynamic question with a main focus on the free energy of transition states.

### Transition State Theory

For the simplest case, let us imagine an isomerization reaction, in which one reactant, **A**, is converted to a product, **B**, by the breaking of one bond and the making of a new one. The progress of any chemical reaction entails the breaking and making of chemical bonds. While the bonds of A and B are stable, there must exist in the middle a **transition state** that is halfway between the two: one bond is about to be broken, and the other is about to be made. The bonds and orbitals are barely decipherable, being in a completely unstable state. It should be clear from this description that **transition states are not any sort of stable intermediates; they are the least stable points on a reaction pathway**. For many reactions the transition state can be so thermodynamically unfavorable that under normal conditions no significant portion of the reactants ever attains the transition state, and the reaction does not occur. For a significant proportion of reactants to attain the transition state, we must reduce the energy penalty of attaining it.

Our first premise is that **the rate of a reaction will be approximately proportional to the concentration of the transition state, ts**. So to increase the *rate* of A→B (or, perhaps, B→A), we need to increase the *concentration* of the transition state. The second premise is that there is, at least approximately, an **ordinary equilibrium between A and the transition state**, t.s.,  $A \rightleftharpoons \text{ts}$ . By reducing the  $\Delta G$  for the  $A \rightleftharpoons \text{ts}$  equilibrium, we increase the **tiny concentration of ts** that is present at any given instant. There are two main approaches to reducing the energy penalty.

The first, favored for instance by anyone who uses fire (which enables the oxidation of fuels), and often by chemists too, is to raise the temperature of the reactant(s). This raises the starting free energy of the system,  $G(\text{start})$ , while  $G(\text{ts})$  remains the same. An alternative, much preferred for obvious reasons in living systems, but a lot in chemistry too, is to use *catalysts*. In *transition state theory*, **enzymes function by binding favorably to transition state(s)**, and thus reducing the free energy of the transition state, to  $G(\text{ts,cat})$ . In other words, the **free energy of the enzyme:ts complex is less than the free energy of ts by itself**. The result is that the free energy that is required to get from A to the transition state is reduced, i.e.  $\Delta G$  for the  $A \rightleftharpoons \text{ts}$  equilibrium is reduced.



### A Numerical Example

Imagine our reaction,  $A \rightarrow B$  (Fig. C1), normally goes at a minuscule rate of just  $10^{-6} \text{ s}^{-1}$  (corresponding to a half-life of A of about 8 days), and that we then introduce a catalyst that reduces the  $\Delta G$  for formation of the transition state **by 10 kcal/mol**. It is then possible to calculate roughly how much the reaction will accelerate in the presence of the catalyst. Gibbs' well-known relationship (following from Boltzmann's work) states that

$$\Delta G = -RT \cdot \ln(K),$$

where  $R$  is the gas constant,  $T$  is the temperature in Kelvin, and  $K$  is the equilibrium constant =  $[ts]/[A]$ . (At room temperature the value of  $RT$  is close to 0.6 kcal/mol.)

**NOTE:** the "equilibrium constant"  $K$  is the chemist's normal meaning of this term,  $[products]/[reactants]$ , i.e. it is like an association constant. This is the reciprocal of the biochemist's more common dissociation constant,  $K_d$ . Always be wary of the plain term "equilibrium constant" when biochemists discuss free energies.

To estimate how much faster the reaction will go, we need to determine the **increase in the concentration of ts in the A $\rightleftharpoons$ ts equilibrium**. We will use  $K_C$  and  $K_{NC}$  to denote the equilibrium constants for the catalyzed and noncatalyzed A $\rightleftharpoons$ ts equilibria, and  $\Delta G_C$  and  $\Delta G_{NC}$  the free energy changes. (**NOTE** that we are talking about A  $\rightleftharpoons$  ts, not the A  $\rightleftharpoons$  B equilibrium.) Watching the minus signs carefully, the difference in  $\Delta G$  when catalyst is present is then

$$\Delta G_{NC} - \Delta G_C = RT[-\ln(K_{NC}) + \ln(K_C)] = RT \cdot \ln(K_C/K_{NC}).$$

Because  $K = [ts]/[A]$ , and  $[A]$  is the same for both, it follows that  $[ts]_C/[ts]_{NC} = K_C/K_{NC}$ . Exponentiating to remove the logarithms, and **using the 10 kcal/mol number for the reduction in  $G_{ts,cat}$**  from my example,

$$\frac{[ts]_C}{[ts]_{NC}} = \frac{K_C}{K_{NC}} = e^{\left(\frac{\Delta G_{NC} - \Delta G_C}{RT}\right)} = e^{(10/0.6)} = 1.73 \times 10^7.$$

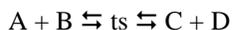
So the catalyst in this example, by reducing the  $\Delta G$  penalty for formation of ts by 10 kcal/mol, **raises the ts concentration about 17 million fold**, and increases the rate by the same fraction—from  $10^{-6} \text{ s}^{-1}$  to  $17 \text{ s}^{-1}$ , i.e. the half life of A is reduced from **8 days to about 40 milliseconds**. Although this was an entirely fake example, it is not out of line for the level of acceleration that enzymes can achieve. Many reactions that enzymes catalyze simply do not occur otherwise. The stability of peptide bonds (i.e. proteins) to hydrolysis in the absence of enzymes is just one example. In this case hydrolysis is thermodynamically favorable ( $\Delta G$  for protein +  $\text{H}_2\text{O} \rightarrow$  peptides is negative), but without the catalyst to enable getting through the transition state the reaction will not occur at any significant rate (fortunately...).

## Reversibility

Before moving on, you must also remember that transition states can break down in either direction, in the previous example either back to A or forward to B. So **what actually makes the net reaction go forward?** (Note that the free energy of B is lower than A in this example, and B is therefore favored.) The answer to this lies with the fact, which we've discussed at length already, that the A  $\rightleftharpoons$  B equilibrium is a dynamic balance of the forward and reverse rates. Because the  $\Delta G$  penalty for B to get up to the transition state is larger than for A, the reverse reaction must be slower.

## Multimolecular Reactions

All we have considered so far is the simplest unimolecular reaction, an isomerization. Changing the concentration of A in that system, although it would increase  $G(\text{start})$ , would also increase  $G(\text{t.s.})$ , because the *proportion* of A able to attain the transition state would remain the same. But unimolecular reactions are rare. In multi-molecular reactions we *can* take advantage of varying reactant concentrations, and this is a very common means of making reactions go, or not, just as the organism, or you in the lab, desire. The key is that the free energies involved [ $G(\text{start})$ ,  $G(\text{end})$ ] describe the *system*, not just the state of the individual molecules. Consider the reaction



and how we might adjust the reaction with respect to B. If, for example, we hold  $[B]$  constant and increase  $[A]$ , this will increase the  $G(\text{start})$  to our benefit, because it will drive a *higher proportion of B* into the transition state, and hence increase the rate of reaction with respect to B. This is what is meant when we talk about "**driving a reaction**".

Because all except unimolecular reactions involve concentrations in this way, the starting free energy of compounds is important information. The **values of  $G$**  that you look up in a reference book are always quoted in *standard terms*. The standard temperature is 25°C, and for compounds in aqueous solution, the standard is a 1 M solution. The two exceptions are

water, which is assumed to be 55.6 M, and hydrogen ions, which are assumed to be  $10^{-7}$  M (pH 7). You then calculate the actual free energy  $G$  of your system from the known reactant concentrations of each component.

### Stabilization of the Transition State

Although there are other important explanations and details, like their role in bringing reactants together in the right steric presentation, **a major role of catalysts is to bind favorably, or stabilize, transition states.** Still, the transition state cannot even approach real stability: we still need to put substantial energy in to form the enzyme:ts complex. That energy comes from favorable interactions with the enzyme. The fact that almost all biological catalysts are proteins is not irrelevant, because proteins, through their multi-functional structures, enable multiple, and 3-dimensionally-correct, interactions with reactants. The total energy of binding available is essentially the sum of the binding energies for each piece of the total interaction. Multiple interactions, as long as they are directed at recognizing the transition state (not the substrate—see below), enable huge reductions in the free energy of the transition state. And as we saw from the example—of the catalytic effect of reducing  $G(\text{ts})$  by just 10 kcal/mol—the effects on reaction rate can be truly enormous.

A corollary of the use of binding energy to stabilize the enzyme:ts complex is that **efficient enzymes will not bind substrates nearly as tightly as they could**, and there is good evidence for this. One example is that inhibitors that mimic transition states almost invariably bind to enzymes considerably more tightly than the substrates do. The converse is often seen in the study of ligand-receptor interactions: because no transition state is involved here, receptors and ligands can be accurately complementary and they usually have substantially higher affinities. Note, however, that these are just general rules.

### Mechanisms

Favorable interactions between enzyme and transition state include all the common ones: mainly ionic (electrostatic) bonds, hydrogen bonds, and hydrophobic interactions. And always remember that  $\Delta G = \Delta H - T\Delta S$ . You know all this already...

- Ionic bonds are found between the four major charged amino acid side-chains: (i) Arg and Lys, which are positively charged, and (ii) Glu and Asp, which are negatively charged. Occasionally charges on ionized His and Cys residues are important too. Charged groups interact with opposite charges on the substrate. The driving forces (i.e. the source of  $\Delta G$ ) are both enthalpy ( $\Delta H$ ) and entropy ( $\Delta S$ ). The latter—a positive  $\Delta S$  term—often arises from the displacement of solvated water by the incoming binding group.

- Hydrogen bonds play major roles, mainly between  $-\text{OH}$ ,  $=\text{O}$ ,  $=\text{NH}$ , and  $-\text{NH}_2$  groups. Here again both enthalpy and entropy play major roles; and once again the entropy component is often caused by the displacement of water, in this case hydrogen-bonded water.

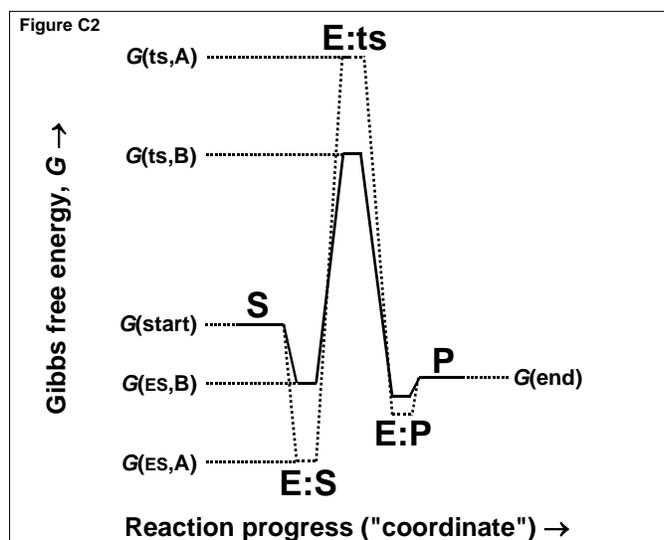
- Hydrophobic interactions are also common. They are almost entirely entropy driven, involving the displacement of ordered water that surrounds a "naked" hydrophobic group in a polar (e.g. aqueous) environment.

Just a single interaction of any of these types can account for a  $\Delta G$  of several kcal/mol, and multiple interactions can easily result in total  $\Delta G$  in the 10's of kcal/mol...almost all of it available for stabilizing the extremely unfavored enzyme:transition state complex. The following discussion is largely due to Alan Fersht, who was early involved in the concept of the efficient enzyme, and its kinetic properties.

Consider a reaction,  $S \rightarrow P$ , that can be catalyzed by two different enzymes, called A and B (Fig. C2).

**Enzyme A** (dotted line) is **inefficient**. It is optimally complementary to the substrate A, and binds it tightly, and there is a substantial negative  $\Delta G$  on binding to form the E:S complex. The enzyme is **not optimally complementary** to the transition state, so the  $\Delta G$  for going from E:S  $\rightarrow$  E:ts, i.e.  $G(\text{E:ts,A}) - G(\text{E:S,A})$ , is very large, and the proportion of E:S able to attain the E:ts transition-state complex will be very small.

**Enzyme B** (solid line) is **optimally complementary** to the transition state. It does not bind substrate particularly well, and the  $\Delta G$  for formation of the E:S complex,  $G(\text{start}) - G(\text{ES,B})$ , is correspondingly small. Because enzyme B binds the transition state better, the  $\Delta G$  for formation of E:ts is much less than for



enzyme A. The result is that the concentration of E:ts is much higher for enzyme B, and the catalyzed reaction will therefore go much faster: **B is much more efficient.**

Efficient enzymes, therefore, increase their catalytic efficiency at the expense of substrate binding. In kinetic terms (see Enzyme Kinetics), this increases  $k_{cat}$  at the expense of  $K_m$  (i.e.  $K_m$  is higher too). As you will see later, taking  $K_m$  far above the substrate concentration is not effective, because the rate will then drop substantially; but certainly it pays an enzyme to evolve to raise its  $K_m$  (and thereby its  $k_{cat}$ ) to about the same order of magnitude as the substrate concentration.

## ENZYME KINETICS

As we have just seen, **enzymes are biological catalysts, which act by binding to substrate(s), acting on them to enable the breaking and making of covalent bonds, and releasing the products.** Enzymes obviously cannot act at a distance, and they therefore must bind their substrates in order to do anything. We thus come to the simplest possible theory of enzyme catalysis:



### Henri, Michaelis, and Menten

Henri in 1902, and Michaelis and Menten in 1913, realized that this model explains something that was already known in the 1890s: that if the concentration of an enzyme substrate is raised, the reaction rate (velocity is a synonym) increases, but only up to a certain point—as we say these days, the rate *saturates* at high substrate concentration. Henri's analysis, essentially duplicated by Michaelis & Menten, was based on the idea that there is an *equilibrium* in the formation of the ES complex, the dissociation constant being called, for the moment,  $K_m$ :

$$ES_{eq} = \frac{k_1 E_{eq} S_{eq}}{k_{-1}} = \frac{E_{eq} S_{eq}}{K_m} . \quad (21)$$

Note all the "eq"s in this expression, to remind you that these are the concentrations of **E, S, and ES at equilibrium.** As shown, the analysis defines the  $E+S \rightleftharpoons ES$  equilibrium by the dissociation constant  $K_m$ . The ES complex can break down either way: to reform E+S, or to go through the transition state and produce products, regenerating the enzyme. **The rate of formation of product(s),  $v$ , is proportional to the concentration of ES**, the rate constant being called  $k_{cat}$ , the catalytic rate constant:

$$v = k_{cat} ES_{eq}, \quad (22)$$

To get an expression for **the rate in terms of E and S**, we therefore need to **calculate the concentration  $ES_{eq}$ .** Calculating this as a function of the initial concentrations of E and S,  $E_0$  and  $S_0$ , is exactly the same as for any other bimolecular interaction (see Eq. 1–7). And just as with the previous analysis, we assume in enzyme kinetics that  $S_0 \gg E_0$ . Thus, getting  $ES_{eq}$  exactly as for Eq. 6,

$$\text{velocity, } v = k_{cat} ES_{eq} = k_{cat} \frac{E_0 S}{K_m + S}, \quad (23)$$

where  $E_0$  is the initial (i.e. total) enzyme concentration. As Henri, Michaelis, and Menten envisioned it, with a true equilibrium  $E+S \rightleftharpoons ES$ ,  $K_m$  is a true dissociation constant. Another way of stating this is (Scheme 1) that  $k_{-1} \gg k_{cat}$ . If this requirement is not met, Eq. 23 does still hold, but  $K_m$  is no longer simply  $k_{-1}/k_1$ . *In practice, the **equilibrium requirement is not met for many enzymes**—perhaps most enzymes acting on their natural substrates—and  $K_m \neq k_{-1}/k_1$ . It is thus not a simple measure of affinity of enzyme and substrate.*

It is relatively simple to extend the M&M sort of analysis to the situation where  $k_{cat} \geq k_{-1}$ , but we are still considering the very simplest model here (Scheme 1): one substrate, one step. Where Michaelis & Menten considered a true equilibrium between E+S and ES, **Briggs and Haldane** considered a less restrictive steady-state assumption for the ES complex.

## Steady States

### An non-enzyme example: regulation of a protein's concentration:

Here I go off on an important tangent that will introduce steady states, and how to analyze them. My favorite is the model of water flowing at a fixed rate into a bucket that has a hole in the bottom, and calculating the depth that is eventually attained, when the rate in equals the rate out, i.e. the *steady-state* depth. But let us consider something a little more relevant: regulating the concentration of a protein in the cytoplasm.

Imagine a particular cytoplasmic protein, called P. Let the rate of synthesis of P, which is controlled mainly by regulation of transcription, be  $R_{\text{formation}}$ , in units of mass or concentration of protein per unit time. Imagine also (certainly a simplification, but acceptable) that the **degradation** of this protein in the cytoplasm is **first-order**, i.e. rate is simply proportional to the concentration of P in the cytoplasm, defined by a first-order rate constant,  $k$ ; so  $R_{\text{degrad}} = k \cdot P_{\text{cyto}}$ . To calculate the cytoplasmic concentration of P we assume that its concentration is due to a dynamic steady state with a balance between synthesis and degradation, i.e.

$$R_{\text{formation}} = R_{\text{degrad}}$$

It follows that at steady state

$$R_{\text{formation}} = k \cdot P_{\text{cyto}}$$

Therefore, the steady state concentration of P in the cytoplasm,

$$P_{\text{cyto}} = R_{\text{formation}}/k.$$

It is obvious that the concentration of P can be controlled equally well by its rate of synthesis *OR* by its rate of degradation.

## Enzymes: Briggs and Haldane

These biochemists in the 1920s were uncomfortable with M&M's assumption about a true  $E+S \rightleftharpoons ES$  equilibrium in enzyme catalysis, and proposed another approach, which is now standard: that **the concentration of ES complex in an enzyme-catalyzed reaction can be assumed to be approximately in a steady state**, where the rate of ES formation closely approximates its rate of disappearance, even though the flux *through* the ES complex,  $S \rightarrow ES \rightarrow P$ , will be very much larger. The assumption, then, is that **at steady state the rates of ES formation and disappearance are equal**:

$$k_1 E \cdot S = k_{-1} ES + k_{\text{cat}} ES = (k_{-1} + k_{\text{cat}}) ES \quad (24)$$

Therefore,

$$ES = \frac{k_1 E \cdot S}{k_{-1} + k_{\text{cat}}}, \quad (25)$$

all concentrations being those at steady state. Compare this with Eq. 21, and you will see that we have replaced Michaelis & Menten's original equilibrium constant,  $K_m = k_{-1}/k_1$ , with a new constant. The derived Eq. 23 still holds, but  $K_m$  is no longer an equilibrium constant; it now includes a  $k_{\text{cat}}$  term:

$$K_m = \frac{k_{-1} + k_{\text{cat}}}{k_1} \quad (26)$$

For the most efficient enzymes,  $k_{\text{cat}} > k_{-1}$ , and **Briggs & Haldane's analysis is closer to the truth**. Under these conditions  $K_m \sim k_{\text{cat}}/k_1$ .

## Initial Rate Requirement

Just as in the derivation of Eq. 6, **the substrate concentration in an individual enzyme experiment must be known**.

Although  $S \gg E$  at the start of just about any reaction you might study, it is much more important to remember that the **substrate concentration must not change significantly** while you measure the velocity of the reaction, so that  $S \sim S_0$  while you measure the rate. As a practical rule, you should try to limit your measurement of rate in any experiment to less than 10% consumption of the substrate. If you are in a bind, 20% might be acceptable, but don't do it often. If you can't manage that, there are occasionally ways of analyzing an entire reaction course (see *First Order Kinetics*, below); but the initial-rate method is preferable, and standard.

Another important reason for measuring initial rates exists **when the reaction being measured is reversible**. Enzymes cannot selectively catalyze a reversible reaction in only one direction, for that would run afoul of the laws of thermodynamics

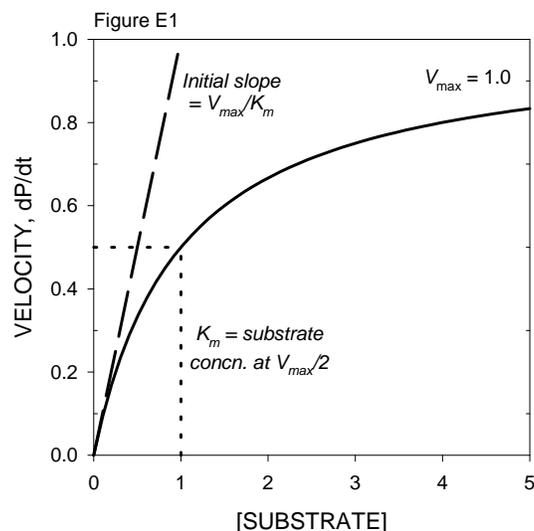
(it would change the  $\Delta G$  for the reaction, which is not allowed). So to measure  $S \rightarrow P$  with no interference from the reverse reaction  $P \rightarrow S$ , we must make sure that  $P$  does not rise to interfering levels while the measurement on  $S$  is being performed. Thus another reason for the measurement of initial rates.

### $K_s$ , $K_m$ , $V_{max}$ , $k_{cat}$ , and $k_{cat}/K_m$

$K_s$ . For the *true* dissociation constant for the  $E+S \rightleftharpoons ES$  equilibrium, we now reserve the term  $K_s$ , i.e.  $K_s = k_{-1}/k_1$ . Ordinary steady-state methods of enzyme kinetic analysis produce no estimates of  $k_1$  or  $k_{-1}$ , and *cannot* be used to determine  $K_s$ .

**$K_m$  is defined empirically, as an apparent dissociation constant, being the concentration of  $S$  that produces half-maximal velocity.**

This follows from Eq. 23: if you set  $S = K_m$ , then  $v = \frac{1}{2}(k_{cat}E_0)$ . It is derived from analysis of the experimental rate data with fixed enzyme and varying substrate concentration (Fig. E1).



**$V_{max}$  is also an empirical term, being the maximum velocity for the enzyme concentration in use.** This occurs with saturating levels of substrate ( $S \gg K_m$ ), so all  $E$  is present as the  $ES$  complex, and  $V_{max} = k_{cat}E_0$ . It is not a directly measured number: **it is derived, along with the  $K_m$ , from analysis of the rate data.**

**$k_{cat}$  is therefore defined as  $V_{max}/E_0$**  (at least for enzymes with only one active site per molecule). If you know the enzyme concentration in your experiment,  $k_{cat}$  is derived directly from  $V_{max}$ . In our simple Scheme 1, it is indeed the first-order rate constant for  $ES \rightarrow P$ . However, for all real-life enzymes, there are more steps involved in getting from  $ES$  to  $P$ , and in this case  **$k_{cat}$ , just like  $K_m$ , is a "lumped" constant.** We will consider an example below.

**$k_{cat}/K_m$ .** A final important parameter that can be obtained from the rate plot is the  $k_{cat}/K_m$  ratio, **even if  $k_{cat}$  and  $K_m$  individually are unknown.**  $V_{max}/K_m$  is the initial slope of the hyperbola, and gives  $k_{cat}/K_m$  directly if  $E_0$  is known.  $k_{cat}/K_m$  is commonly used as a **measure of enzyme efficiency.** (It is an apparent second-order rate constant. Although it does not correspond to any specific step in an enzyme-catalyzed pathway, it must exceed  $k_1$ .) It is also important in sometimes being the *only* parameter that can be determined in experimental attempts at "kinetic characterization". Imagine, for instance, you are studying an enzyme and simply cannot do rate determinations at high substrate concentration...maybe the substrate is ridiculously expensive, or it comes out of solution, or whatever... In that case you will not get rates anywhere close to  $V_{max}$ ; and if you cannot estimate  $V_{max}$  you cannot determine  $K_m$  either (because  $K_m$  is the substrate concentration when  $v = V_{max}/2$ ). You can still, however, determine the  $k_{cat}/K_m$  ratio, and it is a perfectly valid and real description of the activity of the enzyme at low substrate concentrations: **when  $S \ll K_m$ ,  $v = k_{cat}E_0S/K_m$ .**

### $K_m$ , Binding Affinity, and Catalytic Efficiency

It is too often stated by people studying enzymes—for example, comparing the action of an enzyme on a series of substrates—that the  $K_m$  is a measure of the binding affinity of enzyme and substrate. This harks back to Henri and Michaelis & Menten, and for many (maybe most) enzymes acting on their natural substrates it will not be true. As Briggs & Haldane made clear, it is only true (for our simple Scheme 1) if  $k_{-1} \gg k_{cat}$ . It should also be clear from the discussions above about catalysis that the  **$K_m$  is not a measure of enzyme potency** in any shape or form, since—you will remember—efficient enzymes will tend to evolve to *raise* their  $K_m$ 's (at least up to their substrate concentrations) in order to maximize  $k_{cat}$ . For enzyme potency or efficiency comparisons,  **$k_{cat}/K_m$  is a much better parameter to use.**

### First-Order Kinetics, $K_m > S$

As mentioned above, an area of practical importance is the simplification in analysis that occurs when  $K_m > S$ , and

$$v = \frac{dP}{dt} = \frac{k_{cat}E_0S}{K_m}. \quad (27)$$

Because the rate is proportional to  $S$ , the reaction is said to be **"first order in  $S$ ".** In practice this simplification can be used satisfactorily when  $K_m > 3 \times S$  (depending on your tolerance for imperfection—mine is fairly high). The reason I bring it up

is that it is particularly useful when considering **the complete time course of an enzyme-catalyzed reaction**, rather than just the initial rate. Eq. 27 can be integrated to give a time course for P generation *if the reaction is not reversible*:

$$P_t = S_0 \left( 1 - e^{-\frac{k_{cat}}{K_m} E_0 t} \right) \quad (28)$$

Thus a single careful experiment done under conditions where  $S < K_m$  (a full time-course of the production of P with time) can give a value for  $k_{cat}/K_m$ .

## Integrated Michaelis-Menten

What would be really useful would be to be able to integrate the Michaelis-Menten equation (Eq. 23) and produce something equivalent to Eq. 28 that one could simply fit single-time-course data too. There are two problems. First, integration gives what is called an **implicit equation**, with the measured variable (usually  $P_t$ ) appearing on both sides of the = sign:

$$P_t = k_{cat} E \cdot t - K_m \ln \left\{ \frac{S_0}{S_0 - P_t} \right\} \quad (29)$$

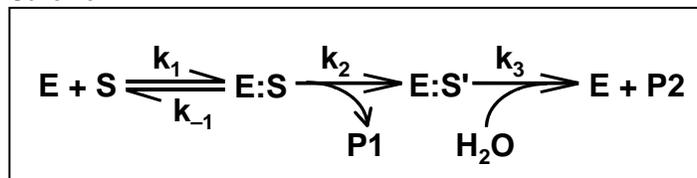
This type of function can be handled by data-handling packages that can fit data to implicit functions. But there is another, more severe, problem: **the method is extremely sensitive to the value of  $S_0$** , and to the assumption that total product formed,  $P_\infty$ , exactly equals  $S_0$ . If there is, for instance, just a little product inhibition or the reaction is reversible to any degree, the method can quickly become extremely unreliable. Not really recommended...

## Multi-Step Reactions

### An Example

No enzyme-catalyzed reaction is actually as simple as Scheme 1. Briggs and Haldane's analysis was specifically based on Scheme 1, and although their approach is applicable to many other reactions, the details about **the precise meaning of  $K_m$  and  $k_{cat}$** —their specific relation to the rate constants in the scheme—vary widely according to the reaction details. A lot of this work dates back to the 1950s and

Scheme 2



1960s, and much was based on the mechanisms of hydrolysis of synthetic substrates by serine proteases, particularly chymotrypsin. Although historical in terms of dates, the general approaches are still valid. Scheme 2 shows a typical hydrolysis reaction. (This basic mechanism is seen in any number of hydrolases, and is by no means limited to proteases.)

For amide and peptide substrates of a protease,  $RCO-NHR'$ , the first E:S complex breaks down to release P1, the amine ( $R'NH_2$ ), with the  $RCO-$  group remaining attached in the second complex, E:S' (the acyl-enzyme complex). Water then comes in and displaces the carboxyl, releasing the second product,  $P2 = RCOOH$ , and the free enzyme. Amides are similar to peptides, and the basic kinetics are the same: **when  $k_3 \gg k_2$  there is no significant buildup of the E:S' intermediate** during the course of the reaction. However, it happens that serine proteases cleave esters too,  $RCO-OR'$ , and these behave very differently. In particular,  $k_2 \gg k_3$ , and E:S' builds up. Thus P1 (the alcohol part of the ester) is released rapidly, while the hydrolysis of E:S' is much slower. (This has been taken advantage of in the design of "enzyme titrants", which react to form an amount of the P1 alcohol that is stoichiometric with the enzyme, but essentially stop at that point because  $k_3$  is so small.)

### Mechanistic Meaning of $K_m$ , $k_{cat}$

Having settled on a likely mechanism, such as Scheme 2, how do we now work out how the empirical kinetic parameters ( $K_m$  and  $k_{cat}$ ) relate to the rate constants involved in the reaction? First, note that when considering hydrolytic reactions, the vast excess of water is always considered as a constant, and not included specifically as a substrate (which it actually is).

Since  $v = k_3 ES'$ , it is clear we need to derive an expression for  $ES'$  at steady state. There are two steady states in this scheme, for the two enzyme + substrate complexes:

$$dES/dt = k_1E \cdot S - k_{-1}ES - k_2ES = 0$$

$$dES'/dt = k_2ES - k_3ES' = 0$$

I will not show the derivation. The important thing here is the meaning of  $k_{cat}$  and  $K_m$  that the derivation enables:

$$K_m = \frac{k_{-1}k_3}{k_1(k_2 + k_3)} \quad (30)$$

and

$$k_{cat} = \frac{k_2k_3}{k_2 + k_3}. \quad (31)$$

These relationships hold only for Scheme 2: other mechanisms—for instance a random-order mechanism instead of an ordered one—will produce different relationships again. **Compare Eq. 30 and 31 with the equivalents for Scheme 1.** For the simpler scheme,  $K_m = (k_{cat} + k_{-1})/k_1$ , and was always *larger* than  $K_s$  (Eq. 26). But in Scheme 2,  $K_m$  is always *smaller* than  $K_s$ .  $k_{cat}$  is also different. In the simpler scheme,  $k_{cat}$  was simply the rate constant for  $ES \rightarrow P$ . Here  $k_{cat}$  is a lumped constant: it is not the rate constant for either  $ES \rightarrow ES'$  or  $ES' \rightarrow P$ , and in fact it must be *smaller* than either of them.

To conclude from just these two examples... In the real world, the **mechanistic meaning of  $K_m$  and  $k_{cat}$  will be a complex function of the reaction mechanism: there is no universal rule.** Thus, for the vast majority of enzymes that have been "kinetically characterized",  $K_m$  and  $k_{cat}$  values are all that are known. (For multi-substrate enzymes,  $K_m$  values for each substrate can generally be obtained with a single  $k_{cat}$ , depending on the particular mechanism: see *Multi-Substrate Reactions*, below.) The key point is that "*characterization*" is largely empirical, not mechanistic. Detailed mechanistic conclusions about multi-step reactions like Scheme 2 can rarely be deduced from  $K_m$  and  $k_{cat}$ .

## Multi-Substrate Reactions

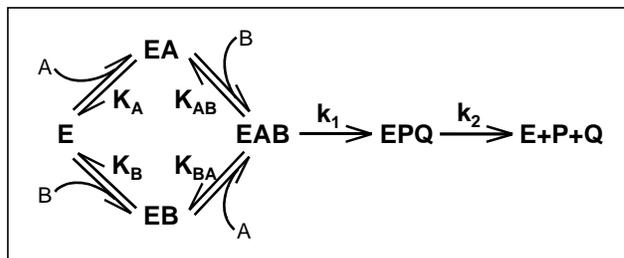
Given the short length of this Kinetics section, I will be brief with multi-substrate reactions. There are many, many different possible mechanisms of enzyme catalysis, and as just discussed, "kinetic analysis" will rarely allow more than initial conclusions about mechanism, *with one major exception*: the binding order of substrates can usually be determined. I will only address the three basic types of substrate addition: 1) random sequential, 2) ordered sequential, and 3) ping-pong. In each case  $K_m$  values can be obtained for each substrate, along with one or more  $k_{cat}$  values. First, the experimental design, which is common to all...

## Experimental

Experimentally, the general design for studying multi-substrate reactions is the same, and it relies on the fact—for any mechanism—that if either **substrate is at a fixed concentration (e.g. A)**, the variation in rate with increasing concentration of the **other(s) (e.g. B) will obey Michaelis-Menten kinetics**, and you will determine an *apparent*  $K_m$  and  $V_{max}$  for B. If you repeat this at varying concentrations of A, you will end up with a collection of data that describes how  $K_{m,app(B)}$  and  $V_{max,app(B)}$  vary with the concentration of A. Since you have in fact done a matrix of experiments at combinations of A and B concentration, you can also plot the data the other way to see how  $K_{m,app(A)}$  and  $V_{max,app(A)}$  vary with B. This type of analysis generates real (non-apparent)  $K_m$  values for each substrate. The specifics of how the apparent parameters for each substrate vary with the concentration of the other substrate are the key to obtaining mechanistic information.

**Scheme 3** encompasses both random and ordered sequential mechanisms. In kinetics, "*sequential*" means those reactions where all reactants combine with the enzyme before any product(s) are released. (Note in contrast that the hydrolysis reaction of Scheme 2 is not "sequential".) In the example, A and B combine with E to form the EAB complex, and the kinetics of its formation depend on the reactant concentrations and their order of addition to the enzyme. Once EAB is formed, the kinetics of product formation will be unaffected by variations in the concentrations of reactants, or the order of their addition to the enzyme, and will depend on the particular mechanism of breakdown of  $EAB \rightarrow EPQ \rightarrow E + P + Q$ .

Scheme 3



## Random Sequential

Here **A and B may combine with E in either order**, so both routes are available to form EAB. While it is common that the binding of one substrate will affect the affinity for the other, the **product of the equilibrium constants in the two pathways must be the same** ( $K_A K_{AB} = K_B K_{BA}$ ), because the  $\Delta G$  for the equilibrium  $E+A+B \rightleftharpoons EAB$  must be independent of the route taken. The experimental approach is the same as for any multi-substrate reaction: determine apparent parameters,  $K_{m,app}$  and  $V_{max,app}$  for each substrate while varying the concentration of the other. For example, at any fixed concentration of A, vary B over a range and measure the velocity in each case. Plot and fit in the normal way to give a  $K_{m,app}$  and  $V_{max,app}$  for B. You then repeat the apparent-parameter determinations for B at varying concentrations of A, producing a whole series of  $K_{m,app}$  and  $V_{max,app}$  estimates for B—in fact a matrix of rate determinations at varying concentrations of both A and B. Four plots can then give you all five parameters ( $K_A$ ,  $K_B$ ,  $K_{AB}$ ,  $K_{BA}$ , and  $V_{max}$ ). They are:

Plot  $V_{max,app}$  obtained for A against the varying concentrations of B used, and *vice versa*.

Plot  $K_{m,app}$  obtained for A against varying concentrations of B used, and *vice versa*.

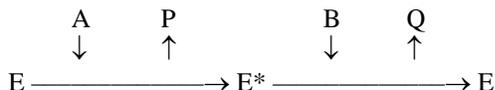
The relationships are given in Segel's Enzyme Kinetics (p.276).

## Ordered Sequential

Here A and B must combine with the enzyme in a specified order, i.e. there is only one order of addition for  $E + A + B \rightarrow EAB$ . This is just a special case of Random Sequential, with one pathway absent. The experimental approach is the same but less exhaustive, because here there are only three parameters to determine.

## Ping-Pong

A ping-pong mechanism means that the first product(s) are released from the enzyme before later substrate(s) bind. They are common in group-transfer reactions, where the first part of the reaction is usually the transfer of the group *to the enzyme* (forming  $E^*$ ). The second substrate then binds and the group is transferred to the second substrate. The general notation is:



The relative independence of the two parts of the process means that there are two  $K_m$ 's and two  $k_{cat}$ 's. The observed kinetics—i.e. the dependency of  $K_{m,app}$  and  $V_{max,app}$  for each substrate on the concentration of the other substrate—are, as you might imagine, quite different from the sequential mechanisms. Again, details are available elsewhere. **The double-reciprocal plots for ping-pong mechanisms are classic examples** of how one can determine mechanism from linearizing analysis: parallel lines are obtained for one substrate at varying concentrations of the other; in contrast to sequential mechanisms, where the double-reciprocal lines intersect in the upper-left quadrant.

## Membrane-Bound Systems

To this point we have considered all enzymes, reactants, and products to be in the solution phase, and all reactions to be described by the ordinary first- and second-order laws of mass action based on 3-dimensional solution concentration. Unfortunately biochemistry is replete with processes that occur on surfaces like membranes. (Other "surfaces" exist, like chromatin DNA, super-macromolecular complexes, etc., but deciding whether they are surfaces or solutions is up to the beholder.) The ones I address briefly here are genuinely 2-dimensional, and membranes are the best example. Reactions may occur where both enzyme and substrate are membrane-bound, i.e. purely in two dimensions; or they may occur where you have membrane-bound enzyme and a solution-phase substrate, producing mixed dimensionality. The detailed kinetics of the mixed systems are very hard to study or analyze without resorting to techniques that allow some assumption of solution-type kinetics, and the measurement of reactant concentrations in pseudo-solution terms. This is in fact the normal approach (see Experimental, below).

Planar concentrations must of course be used in 2-dimensional systems. A useful planar concentration unit is  $\text{pmol}/\text{cm}^2$ , which happens to be of the right order of magnitude to give reasonable numbers (in biochemistry at least) without resort to very large powers of ten, e.g. a membrane containing a near-saturating level of a typical protein may contain of the order of 1-10  $\text{pmol}/\text{cm}^2$ , and a reasonable cell receptor density will likely be in the range of 0.01 – 1  $\text{pmol}/\text{cm}^2$ . For both enzyme and substrate on a membrane, a **Michaelis-Menten type of analysis is perfectly feasible**, with enzyme concentration, substrate concentration, and  $K_m$  all specified in **planar units**. With these numbers, one can calculate the ES concentration in the same



## Competitive

**Competitive inhibitors compete with a substrate.** They often bind to some part of the enzyme's substrate-binding site. For example, aromatic amidines like benzamidine inhibit trypsin-like proteases because they closely mimic an Arg or Lys side chain, and occupy the enzyme's primary substrate-binding pocket. In competitive inhibition **only E combines with I** ( $K_i$  exists, Scheme 4). The EI complex cannot combine with S; nor can the ES complex combine with I ( $K'_m$  and  $K'_i$  do not exist). The balance between the two equilibria (enzyme binding to S or I), and thus the extent of inhibition, **depends on the concentrations of both I relative to  $K_i$  and S relative to  $K_m$ .** The effect of a competitive inhibitor is to draw the  $E+S \rightleftharpoons ES$  equilibrium to the left, *reducing* the concentration of ES; and the *observed result* is an apparent reduction in affinity of enzyme and substrate, i.e.  $K_m$  increases:

$$v = \frac{E_0 k_{cat} S}{S + K_m \left( \frac{K_i + I}{K_i} \right)} \quad (33)$$

You will note that  $V_{max}(k_{cat})$  stays the same, i.e. **high concentrations of substrate can overcome inhibition.** Only when  $S \ll K_m$  will the extent of inhibition be decided solely by  $(K_i + I)/K_i$ .

## Uncompetitive

This is the other extreme of the inhibitory mechanisms in Scheme 4, where **I binds only to the ES complex, not to E**; i.e.  $K_m$  and  $K'_i$  exist, but  $K'_m$  and  $K_i$  do not. This mechanism is rare in simple systems, although it is sometimes seen in multi-substrate systems. The effect of I is to draw the  $E+S \rightleftharpoons ES$  equilibrium to the right, and increase the apparent affinity of E and S. Thus  $K_m$  decreases. At the same time, however, inhibitor is taking ES out of the system, and  $k_{cat}$  is therefore also reduced, by the same factor:

$$K_{m,app} = \frac{K_m K_i}{K_i + I}, \text{ and } k_{cat,app} = \frac{k_{cat} K_i}{K_i + I}.$$

Once again, **because  $K_m$  is involved the extent of inhibition will depend on S relative to  $K_m$ .**

## Mixed, Noncompetitive

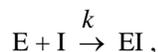
In these two cases, **inhibitor can combine with both E and ES**, i.e. both EI and EIS complexes exist. If the affinities of the inhibitor for E and ES are different, i.e.  $K_i \neq K'_i$ , then inhibition is Mixed. **This is a common mode of inhibition.** Depending on the particular values of  $K_i$  and  $K'_i$ , various extents of increase in  $K_m$  and reduction in  $k_{cat}$  will be observed. Depending on the details, the effect of S on the extent of inhibition will also be variable.

Noncompetitive inhibition is a rare special case in which **I binds with equal affinity to E and ES**, i.e.  $K_i = K'_i$ . Because E and ES are reduced by the same extent in the presence of inhibitor, a noncompetitive inhibitor acts just as if the total enzyme level had been reduced. Thus  $k_{cat}$  is reduced by  $K_i/(K_i+I)$ , but  $K_m$  remains unchanged. **This rare case is the only one where the extent of inhibition does not depend on substrate concentration.**

## IRREVERSIBLE AND TIGHT-BINDING INHIBITORS: RATES

The dividing line between reversible and irreversible is fuzzy. It is really a question of how *you* are going to describe the inhibitory properties. If an inhibitor acts rapidly, and you always describe it in equilibrium terms—you use  $K_i$  to define it—then you are considering it as reversible. On the other hand, if the **rate of action** of an inhibitor is of prime importance, then you are handling it more as an irreversible inhibitor. As you might expect, there are situations, where both the rate of inhibition, and the final extent, are both important. Some inhibitors are genuinely irreversible, but most biological inhibitors—protease inhibitors, for example—are not.

The very simplest mechanism of inhibition is



and the rate,  $dE/dt = -k \cdot I \cdot E$ . Usually inhibitor will be in considerable excess over enzyme, so that  $k \cdot I$  is a constant, and is an apparent first-order rate constant. Therefore the inhibition of E when  $I \gg E$  is characterized by what we call **pseudo-first-order kinetics**. Since  $k \cdot I$  is constant, the rate equation can be integrated to give the time course of decay of E:

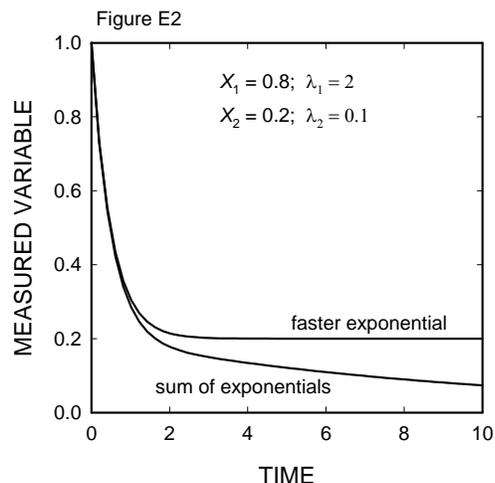
$$E_t = E_0 \cdot e^{-kI \cdot t}$$

## TWO EXPONENTIALS: INHIBITORS AND OTHER CASES

Many processes in all sorts of systems involve the removal or disappearance of things that are not described by simple first-order processes. I will just use two examples, below. First, though, let us consider the kinetics that might be observed. The key phrase is *a sum of exponentials*. The basic equation is as follows (using X to represent any component that is disappearing with time):

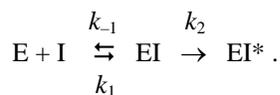
$$X_t = X_1 \cdot e^{-\lambda_1 t} + X_2 \cdot e^{-\lambda_2 t} + \dots \quad (34)$$

Let us consider the two-exponential version. The time course is described by 1) two *amplitudes*, called  $X_1$  and  $X_2$ , and two *rate constants*, which I will call  $\lambda_1$  and  $\lambda_2$ . [Practitioners of **relaxation kinetics** (imagine that discipline!) often deal in something called a relaxation time, denoted by the Greek letter *tau*, which is the inverse of a first-order rate,  $\tau = 1/\lambda$ .] Fig. E2 shows a typical example of the sum of two exponentials (lower curve), plus the faster (and in this case larger) exponential separated out. The difference between the two lines is due to the slower exponential.



### Inhibitors

For many tight-binding inhibitors, things are not as simple as first-order because the mechanism involves two stages, with a rapid initial stage followed by slower formation of the final stable EI complex, here called EI\*:



The time course of inhibition (i.e. disappearance of E) is characterized by two exponentials, Eq. 34 (Fig. E2), and it is usually quite easy to get disappearance data to fit the function. The hard part is devising experiments (and measuring the disappearance kinetics for each) that allow you to derive the three rate constants. As you might guess, varying  $I$  and getting values for  $X_1$ ,  $X_2$ ,  $\lambda_1$ , and  $\lambda_2$  produces useful information, because only one reaction (the formation of EI) is dependent on  $I$ . Increasing  $I$  should have two effects: 1) the first amplitude,  $X_1$ , should increase, because more E is driven into I in the  $E+I \rightleftharpoons EI$  steady state; and 2) the first rate,  $\lambda_1$ , should increase because of the effect of I on  $E+I \rightarrow EI$ . Thus  $k_1$  can usually be derived.

If  $k_{-1} \ll k_2$ , then  $k_2$  can often also be derived, because if  $k_1$  has been determined, then the values of the relative amplitudes ( $X_1/X_2$ ) with varying  $I$  will give you information about ratio of  $k_1/k_2$ . If  $k_{-1} \geq k_2$ , deriving their values can be very hard or impossible.

### Clearance kinetics; multiple pools

These get a mention here because the kinetics are so similar. Things that disappear from systems come up all over the physiological, and nonphysiological, worlds. Imagine that you are studying, for instance, a drug that is taken up by hepatocytes in the liver. You find the clearance data after injection do not fit a single exponential, but they do fit Eq. 34. One possible explanation is that the drug is in two different pools, e.g. one in the blood itself, and another perhaps extravascular. In this case the results of fitting to Eq. 34 are more directly meaningful than in the Inhibitors case above. In the simplest scenario  $X_1$  and  $X_2$  are direct measures of the sizes of the two pools of drug;  $\lambda_1$  is the first-order rate due to uptake by the liver receptor; and  $\lambda_2$  reflects the slower transport rate of the drug from the second pool into the blood.

## COMPETITION WITH INHIBITORS

### By Inhibitors

The reaction of an enzyme with an irreversible inhibitor can also be subject to inhibition. It is sometimes common to be dealing with the two types of inhibitor present together: reversible, and irreversible. It is useful to consider what happens in this case. It is especially relevant to the use of cocktails of protease inhibitors to prevent proteolytic degradation during cell fractionation and protein purification. It is also of course applicable to the inhibition of other enzymes.

Imagine you have a mixture containing active enzyme(s) that you want to inhibit, permanently, in a crude cell homogenate. You add both a competitive inhibitor, I1, and an irreversible one, I2. Assume that both are in large excess over enzyme. The rate of inhibition by I2,  $dE/dt = -k \cdot I2 \cdot E$ . However, the free enzyme concentration,  $E$ , is reduced by the presence of I1 and the formation of the E:I1 complex. The fractional reduction in the rate of inhibition is therefore  $K_i/(K_i + I1)$ . Thus the rate of inhibition by I2,

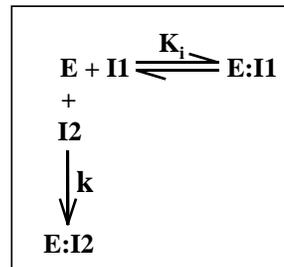
$$-\frac{dE}{dt} = E \cdot I2 \cdot k \frac{K_i}{K_i + I1} \quad (35)$$

You have thus reduced the rate of permanent inactivation of E by including the competitive inhibitor.

### By Substrates

Enzyme substrates will also affect *rates* of inhibition, because they too compete for enzyme, just like I1 in Scheme 5. A common experimental method for measuring rates of inhibition is to include an enzyme substrate along with enzyme and inhibitor, and measure the continuous reduction in rate of product formation as inhibition proceeds. Because substrate and inhibitor are competing for the enzyme, the rate of inhibition, just as above, is reduced by the factor  $K_m/(K_m + S)$ .

Scheme 5



## APPENDIX

### Functions and Shapes

(In each section the species shown in the figure is shown **bold**.)

#### HYPERBOLAS

Hyperbolas describe equilibrium situations, such as the variation in the equilibrium level of a bimolecular complex as a function of the initial concentrations of reactants. Hyperbolic data are never time data. There are two major classes.

(1) *Formation/saturating*, e.g.

receptor + ligand,  $\mathbf{R+L} \rightleftharpoons \mathbf{RL}$

enzyme + substrate,  $\mathbf{E+S} \rightleftharpoons \mathbf{ES}$

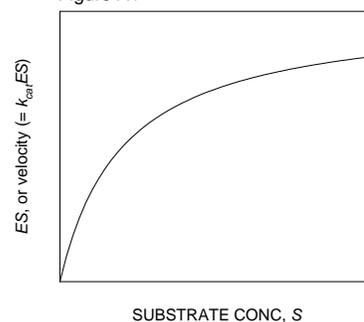
(velocity =  $k_{cat}ES$ )

enzyme + inhibitor,  $\mathbf{E+I} \rightleftharpoons \mathbf{EI}$

example:

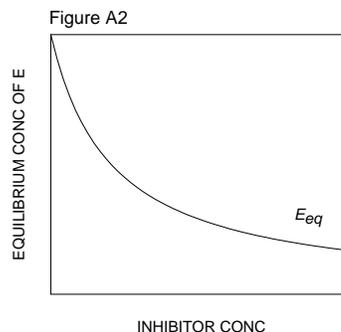
$$ES = \frac{E_0 S}{K_M + S} \quad (A1)$$

Figure A1



2) Depletion/inhibitory/competitive, e.g.  
 enzyme + substrate,  $E+S \rightleftharpoons ES$   
 enzyme + inhibitor,  $E+I \rightleftharpoons EI$   
 receptor + ligand,  $R+L \rightleftharpoons RL$   
 example:

$$E = \frac{E_0 K_i}{K_i + I} \quad (\text{A2})$$



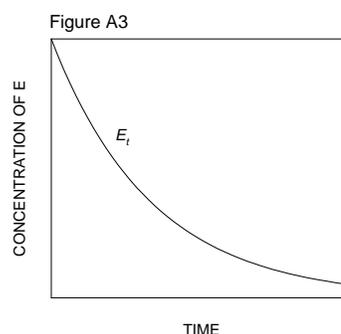
## EXPONENTIALS

**Exponentials describe time-dependent processes, such as how fast a bimolecular complex is formed upon mixing two reactants, or the kinetics of dissociation of a complex. Exponential data are never found in the analysis of equilibrium situations. Two classes, again.**

1) Decay/dissociation, e.g.  
 enzyme + inhibitor  $\rightarrow EI$   
 enzyme:substrate  $\rightarrow$  enzyme + product  
 receptor:ligand  $\rightarrow$  receptor + ligand  
 example:

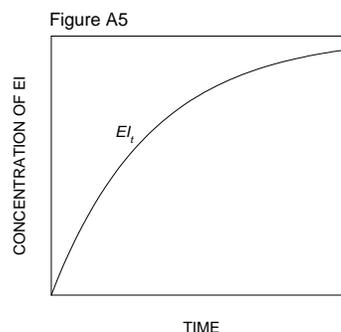
$$E_t = E_0 e^{-kt} \quad (\text{A3})$$

**Half-life,  $t_{1/2} = -\ln(1/2)/k = 0.693/k$**



2) Formation/combination, e.g.  
 enzyme + inhibitor  $\rightarrow EI$   
 receptor + ligand  $\rightarrow RL$   
 example:

$$RL_t = RL_{\max}(1 - e^{-kt}) \quad (\text{A4})$$



## Fitting Data to Functions

### Linearizing

Until the late 1980s investigators did not usually have the means to fit data to nonlinear functions. This is now common in a number of data-handling packages (and a few add-ons for spreadsheets) for PCs. Before that time the functions to which data could be fitted were those that could be *linearized* and then fitted to a straight line by linear regression. Straight lines are easily fitted with a calculator or spreadsheet without having to resort to iterative methods. However, linearizing transforms often mess up the statistics badly. The standard algorithm for nonlinear fitting is due to Marquardt, and it produces error estimates on each fitted parameter and a measure of the goodness of fit, called the reduced  $\chi^2$ .

**It is best in general to fit your data by nonlinear regression to the appropriate function directly. If you then want to visualize the results better with a linear plot, you can linearize both the fitted function and the data.**

**Eq. A1** is commonly linearized as a **double-reciprocal**, or—in enzyme kinetics—**Lineweaver-Burke** plot, e.g.  $1/v$  is plotted against  $1/S$ , giving a Y-axis intercept of  $1/V_{\max}$ , an X-axis intercept of  $-1/K_m$ , and a slope of  $K_m/V_{\max}$ :

$$\frac{1}{v} = \frac{1}{V_{\max}} + \frac{K_m}{V_{\max}} \cdot \frac{1}{S}$$

**This transformation is by far the worst** with respect to statistical errors produced by the transform: the two reciprocals give *maximum* statistical weight to the *smallest* data values, and these are usually the *least* accurately measured. The weighting error at the end of the line (the smallest  $S$  and  $v$  values) is approximately to the 4th power of  $1/S$ , e.g. if you halve  $S$ , the weighting error of  $1/S$  goes up about 16-fold.

Receptor-ligand people used to commonly use another linearizing transform of Eq. A1 called the **Scatchard plot**. For  $R+L \rightleftharpoons RL$ , the (bound ligand)/(free ligand) ratio, i.e.  $RL/L$ , is plotted against the bound ligand ( $RL$ ). The slope is  $-1/K_d$ , and the **X-axis intercept** is the concentration of receptor sites ( $R_0$ ):

$$\frac{RL}{L} = \frac{R_0}{K_d} - \frac{RL}{K_d}$$

Scatchard plots are very nonintuitive plots for people like me who don't use them often, but they do not have nearly the statistical weighting problems of the double-reciprocal plot. Like the double-reciprocal plot, they are useful for *visualizing* things. They have a close relative called the **Eadie plot**, where the axes are simply swapped: the Y axis intercept is  $R_0$ , and the slope becomes  $-K_d$ . I find the Eadie to be a little easier to understand.

**Eq. A2** is linearized as a **Dixon plot**, e.g.  $1/E$  (or  $1/v$ ) is plotted against  $I$ , giving a Y-axis intercept of  $1/E_0$  and a slope of  $1/E_0K_i$ . Once again the reciprocal transform gives greatest weight to the smallest data values.

$$\frac{1}{E} = \frac{1}{E_0} + \frac{I}{E_0K_i}$$

**Eq. A3** is linearized by taking natural logarithms, so that the intercept is  $\ln(E_0)$  and the slope is  $-k$ :

$$\ln(E_t) = \ln(E_0) - kt$$

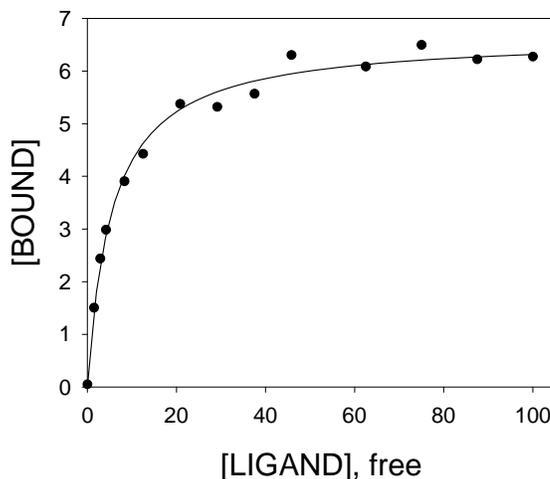
**Eq. A4** cannot be linearized unless the asymptote ( $RL_{\max}$  in the example) is accurately known, in which case you can subtract the data from  $RL_{\max}$ , log transform, and use Eq. A3. Nonlinear regression is much better.

## Which Model Best Fits Your Data?

### Occam and his Razor: Goodness of Fit

How do you know your model is the best one for the data? There are easy, rigorous methods of determining which model is best applicable to your data. The fundamental rule is from William of Occam, a medieval English monk (d. 1350), and is called **Occam's Razor**. He said in essence said that the **best** explanation of an event or an observation must always be the **simplest explanation** that fits. As a rule for scientists, more complex explanations may sometimes turn out to be true, but they **must** be based on further observation that rules out the simpler explanation. Imagine, for example, that you are doing a receptor-ligand study. You suspect that there are two receptors for your ligand, and this obviously produces a bias in how you analyze the data. In a beautiful experiment you obtain these data.

You fit them directly to a 2-site model—on the basis of your suspicion rather than the data—and obtain the best-fit line shown in the figure. The estimates for the receptor concentrations and  $K_d$  values are:



2-site model ( $R1$  and  $R2$  denote the respective receptor concentrations, i.e. max  $L$  bound to each one):

$$\begin{aligned} R1 &= 2.73; & K_d &= 14.3 \\ R2 &= 4.17 & K_d &= 3.26 \\ \text{Reduced } \chi^2 &= 0.0395 \end{aligned}$$

Reduced  $\chi^2$  (chi-squared) is calculated from the distances that the points fall from the best-fit line, and it also takes into account the complexity of the function that you are fitting the data to (i.e. the number of parameters that you are determining: here it is 4). For experiments where you have multiple measurements for each data point  $\chi^2$  can be used to calculate a probability that the data fit the function; but such experiments are unusual (e.g. for the experiment shown, getting 5 measurements at each ligand concentration would mean 70 measurements altogether, for a single line). More usually, **reduced  $\chi^2$  is used to compare the goodness of fit** when a data set is fitted to different functions. Anyway, you think from the beautiful fit that this *must* be the right model, and start writing the paper... Wrong move...

Although the program managed to fit the data to the 2-site model, at **there is no evidence that this is the best** (i.e. most appropriate, or defensible) model for these data. To provide evidence, you must fit the data to all simpler models, and compare the goodness of fit for each. For binding data like these, the simpler models are (i) 1 binding site, with no "nonspecific" (i.e. linear) binding, and (ii) 1 binding site, plus "nonspecific" binding. You re-fit the data to both. The actual lines for each fit were just about indistinguishable from the one shown on the plot—all gorgeous!—and the results of fitting were:

1-site model:

$$\begin{aligned} R &= 6.66; & K_d &= 5.50 \\ \text{Reduced } \chi^2 &= 0.0381 \end{aligned}$$

1-site model + "nonspecific" (i.e. linear) term:

$$\begin{aligned} R &= 6.38; & K_d &= 5.01; & \text{"Nonspecific" constant} &= 0.0034 \\ \text{Reduced } \chi^2 &= 0.0388 \end{aligned}$$

The reduced- $\chi^2$  numbers for the three fits show that the **best** model is the simplest one, of 1 binding site with no "nonspecific" binding. In fact the 2-site model is the worst of all three by this objective measure. You might argue that the reduced- $\chi^2$  numbers are really close, and therefore that the data could still support your 2-site model. Sorry, but no: the trouble with this argument is that **it fails Occam's test for simplicity**.

[In passing, there's another thing of interest here, which came (unplanned) from the way I generated the fake data shown in the figure. The original 14 data points were actually a perfect depiction of a 2-site model, with  $K_d$  values differing by 5-fold, but then I added a very small amount of statistical noise to the Bound values measured (SD of the noise was only  $\pm 0.2$ ). This tiny bit of random noise, equivalent to very small experimental error, turned out to be sufficient to wipe out all evidence that a 2-site model was involved.]

Programs that report error estimates on the values of the fitted parameters also show up problems with fitting data to over-complex models: as the number of parameters rises, they are determined much less accurately.