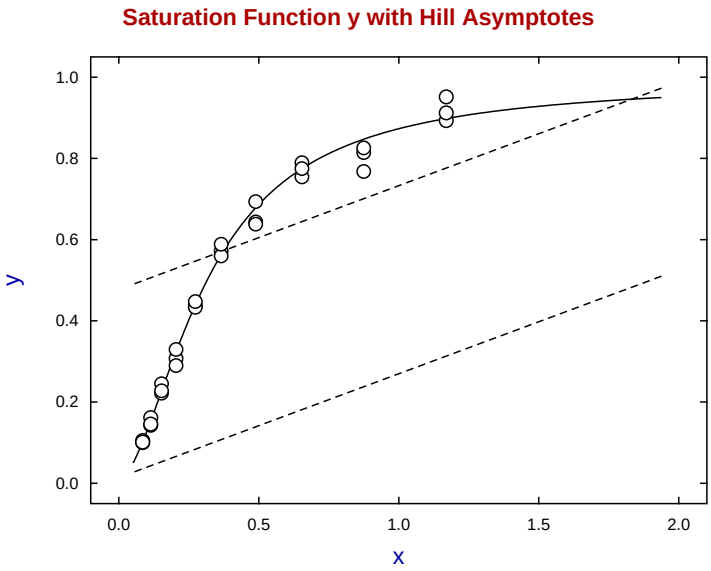
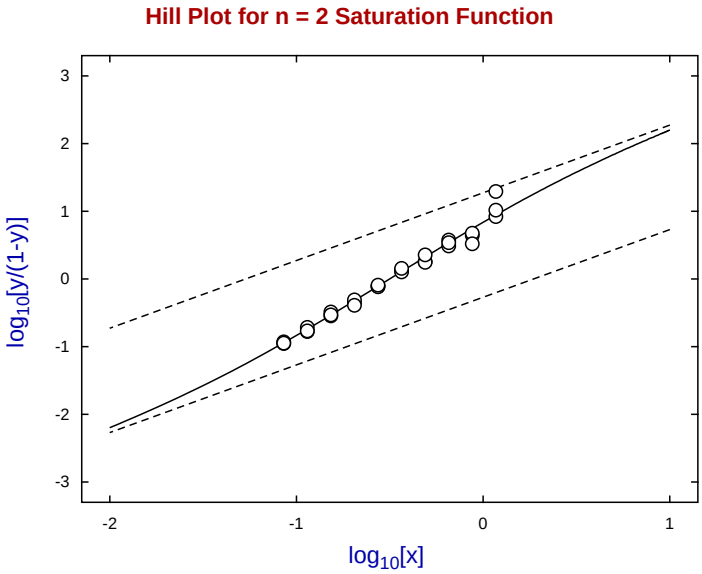
The Hill plot is used to display cooperative effects in ligand binding to a protein or receptor with at least two linked sites. Sites with no linkage will show one-site binding isotherms and so, to explain how to create Hill plots, use program **sffit** to fit a two site model to the test file `sffit.tf4`. After fitting, a cooperativity analysis interface is presented and, choosing the Hill plot, generates the next plot with data, best-fit curve, and extreme points needed to plot the asymptotes in Hill space.



Choosing a Hill plot with asymptote $A = 1$ the next plot will be displayed, and this is followed by an introduction to the principles of cooperativity analysis needed to interpret Hill plots..



Theory

Ligand binding theory will be presented under the following headings.

1. Historical introduction
2. Binding polynomials
3. Definition of cooperativity
4. Factorability of the binding polynomial
5. Statistical interpretation of saturation functions
6. Cooperativity analysis

Historical Introduction

In 1910 Hill [1] proposed that the sigmoid binding curve for oxygen binding to haemoglobin could be analyzed in terms of the binding of n ligands in one step with no appreciable intermediates, i.e. the mass action description

$$Hb + nX \rightleftharpoons HbX_n.$$

This leads to the Hill equation describing the fractional saturation y as a function of concentration x, and the Hill plot of $\log[y/(1-y)]$ as a function of $\log x$ as follows

$$y = \frac{Kx^n}{1 + Kx^n}$$

$$\log\left(\frac{y}{1-y}\right) = n \log x + \log K.$$

It is now realized that the Hill equation is simply an empirical equation that is at best a poor approximation to any real binding situation since:

1. it is only an appropriate representation for a one-site binding process, i.e. for $n = 1$;
2. when $n < 1$ it has an infinite slope at the origin and cannot model any realistic binding situation;
3. when $n > 1$ it has zero slope at the origin and cannot model any realistic binding situation;
4. when n is not a positive integer it is pure nonsense; and
5. using it to discuss the effect of cooperativity on graphical features such as sigmoidicity in the $y(x)$ curve, or convexity in Lineweaver-Burke or Scatchard space, has resulted in considerable confusion.

Of course, before the days of computers and nonlinear regression, fitting a straight line to a Hill plot to get a non-integer value for the estimated slope was all that could be done, and this non-integer value was correctly taken to mean that this was a result of the model being incorrect.

Nowadays no one would dream of discussing cooperative binding in terms of the Hill equation or fitting a straight line to a Hill plot but, by a serendipitous coincidence, it turns out that the variable slope of the curve obtained by transforming a saturation curve into Hill space still provides an unambiguous definition of the sign and magnitude of cooperativity that has got nothing at all to do with the Hill equation. That is because, to use receptor terminology,

$$\frac{y}{1-y} = \frac{[\text{Bound}]}{[\text{Free}]}.$$

Binding polynomials and their Hessians

In 1925 Adair [2] improved the description of binding isotherms by defining binding constants for the individual binding events, and later it came to be appreciated that these have to be normalized by statistical factors in order to discuss the affinity of receptor for ligand in adjacent binding events. In 1967 Wyman [3] rationalized the situation by pointing out that, for a non-aggregating macromolecule with n binding sites and only one ligand x varied, there would be binding polynomial which would act like a partition function in that successive terms of degree i in the polynomial are proportional to the amount of macromolecule with i ligands attached.

So now the binding of ligands to receptors can be defined for all possible cooperative binding schemes in terms of a binding polynomial $p(x)$ in the free ligand activity x, as follows

$$\begin{aligned} p(x) &= 1 + K_1 x + K_2 x^2 + \cdots + K_n x^n \\ &= 1 + A_1 x + A_1 A_2 x^2 + \cdots + \prod_{i=1}^{n} A_i x^n \\ &= 1 + \binom{n}{1} B_1 x + \binom{n}{2} B_1 B_2 x^2 + \cdots + \binom{n}{n} \prod_{i=1}^{n} B_i x^n, \end{aligned}$$

where the only difference between these alternative expressions concerns the meaning and interpretation of the binding constants. The fractional saturation is just the scaled derivative of the log of the polynomial with respect to $\log(x)$, and an important auxiliary function is $h(x)$, the Hessian of the binding polynomial defined as follows

$$\begin{aligned} y(x) &= \left(\frac{1}{n}\right) \frac{d \log p(x)}{d \log x} \\ &= \left(\frac{1}{n}\right) \frac{x p'(x)}{p(x)} \\ h(x) &= n p p'' - (n-1) p'^2. \end{aligned}$$

Definition of cooperativity

Given a binding polynomial of degree n there are $n-1$ cooperativity coefficients c_i defined as

$$c_i = B_{i+1} - B_i \text{ for } i = 1, 2, \ldots, n-1,$$

or alternatively as $\log(B_{i+1}/B_i)$, and the interpretation of these is perfectly clear: in a situation where $c_i > 0$ the macromolecule has greater affinity for binding the $i+1$th ligand after the ith ligand has been bound and it is perfectly reasonable to describe this as mechanistic positive cooperativity. Hence every binding situation for n ligands can be summarized by a succession of $n-1$ signs and it might be thought that during the actual saturation of macromolecule with ligand there would be a succession of phases with possibly differing cooperativity. For instance, the sequence $+ - +$ might be supposed to give a saturation curve with positive, then negative, then positive cooperativity. Unfortunately the cooperativity coefficients cannot be interpreted in this way and they are not a unique indicator of the sign and magnitude of the type of cooperativity exhibited during the saturation process. The reason for this is simply that binding does not occur in a succession of isolated steps and at every stage for $0 < x < \infty$ every species that is possible is present, that is no ligands bound, one ligand bound, two ligands bound, etc. up to n ligands bound.

At every point in the range $0 < x < \infty$ there is a one site binding curve y_{app} with a uniquely defined apparent binding constant K_{app} according to the scheme

$$[\text{Free sites}] + X \rightleftharpoons [\text{Occupied sites}]$$

that is

$$y_{app}(x) = \frac{K_{app}x}{1 + K_{app}x}.$$

Surely all would agree that the sign and magnitude of cooperativity at that point in the saturation curve would depend on whether K_{app} is increasing or decreasing as a function of x. It turns out that

$$K_{app} = \frac{p'(x)}{np(x) - xp'(x)} \text{ and}$$
$$\frac{dK_{app}}{dx} = \frac{h(x)}{(np(x) - xp'(x))^2}$$

so that increasing affinity (i.e. positive cooperativity) requires $h(x) > 0$, decreasing affinity (i.e. negative cooperativity) requires $h(x) < 0$ while at a point where $h(x) = 0$ cooperativity changes sign. Bardsley and Wyman [4] emphasized that the magnitude of the Hill slope with respect to 1 is the unambiguous indicator of cooperativity which also depends on the sign of the Hessian as follows

$$\frac{d\log[y/(1-y)]}{d\log x} = 1 + \frac{xh(x)}{p'(x)(np(x) - xp'(x))}.$$

and Wood and Bardsley [5] proved that the Hessian can have at most $n - 2$ positive zeros.

Zeros of the binding polynomial

If the n zeros of the binding polynomial are α_i then the fractional saturation y can be expressed as

$$y = \left(\frac{x}{n}\right)\sum_{i=1}^{n}\frac{1}{x - \alpha_i},$$

but further discussion depends on the nature of the zeros.

First observe that, for a set of m groups of receptors, each with n_i independent binding sites and binding constant k_i, then the zeros are all real and

$$p(x) = \prod_{i=1}^{m}(1 + k_i x)^{n_i},$$
$$\text{and } y = \frac{1}{\sum_{i=1}^{m} n_i}\sum_{i=1}^{m}\frac{n_i k_i x}{1 + k_i x},$$

so y is just the sum of simple binding curves, giving concave down double reciprocal plots, etc.

Actually Bardsley et al [6] and [7] proved that, if a binding polynomial factorizes into m polynomials p_i with positive coefficients according to

$$p(x) = p_1(x)p_2(x)\ldots p_m(x)$$

then the Hill plot slope cannot exceed that of the Hill plot slope for any of the individual factors. As a binding polynomial can always be factorized into a product of linear factors with real negative zeros and complex conjugate pairs forming quadratic factors it might be supposed that the Hill slope can never exceed two. However, if a binding polynomial of degree > 2 has complex conjugate zeros, the Hill slope may exceed two and there may be evidence of strong positive cooperativity. That is why Hill plot slopes up to a maximum of the degree of the binding polynomial can be achieved if there are quadratic factors with negative coefficients, corresponding to a group of at least three linked binding sites.

For instance, the binding polynomial for a four site Monod-Wyman-Changeux model is

$$p(\alpha) = \frac{1}{1 + L}\left((1 + \alpha)^4 + L(1 + c\alpha)^n\right)$$

and this can factorize into the form

$$q(x) = (1 + a_1 x + b_1 x^2)(1 - a_2 x + b_2 x^2)$$

with $a_1 > 0, a_2 > 0, b_1 > 0, b_2 > 0$ under certain constraints so that the meaningless quadratic factor with a negative term allows Hill slopes greater than two.

Edelstein and Bardsley [8] subsequently explored the relationship between the Hill slope at half-saturation and the Hessian of the binding polynomial.

Statistical interpretation of saturation functions

The species fractional populations s_i which are defined for $i = 0, 1, \ldots, n$ as

$$s_i = \frac{K_i x^i}{K_0 + K_1 x + K_2 x^2 + \cdots + K_n x^n}$$

with $K_0 = 1$, are interpreted as the proportions of the receptors in the various states of ligation as a function of ligand activity. The species fractions defined as $y_i = i s_i / n$ for $i = 1, 2, \ldots, n$ are the contributions of the species to the overall saturation. Note that

$$\sum_{i=0}^{n} s_i = 1, \text{, while}$$
$$\sum_{i=1}^{n} y_i = (1/n) d \log p / d \log x.$$

Such expressions are very useful when analyzing cooperative ligand binding data and they can be generated from the best fit binding polynomial after fitting binding curves with program **sffit**, or by interactive input of binding constants into program **simstat**. At the same time other important analytical results like factors of the Hessian and minimax Hill slope are also calculated.

The species fractional populations can be also used in a probability model to interpret ligand binding in several interesting ways. For this purpose, consider a random variable U representing the probability of a receptor existing in a state with i ligands bound. Then the the probability mass function, expected values and variance are

$$\begin{aligned}
P(U = i) &= s_i \ (i = 0, 1, 2, \ldots, n), \\
E(U) &= \sum_{i=0}^{n} i s_i, \\
E(U^2) &= \sum_{i=0}^{n} i^2 s_i, \\
V(U) &= E(U^2) - [E(U)]^2 \\
&= x \left(\frac{p'(x) + x p''(x)}{p(x)} \right) - \left(\frac{x p'(x)}{p(x)} \right)^2 \\
&= n \frac{dy}{d \log x},
\end{aligned}$$

as fractional saturation y is $E(U)/n$. In other words, the slope of a semi-log plot of fractional saturation data indicates the variance of the number of occupied sites, namely; all unoccupied when $x = 0$, distribution with variance increasing as a function of x up to the maximum semi-log plot slope, then finally approaching all sites occupied as x tends to infinity. You can input binding constants into the statistical calculations procedure to see how they are mapped into all spaces, cooperativity coefficients are calculated, zeros of the binding

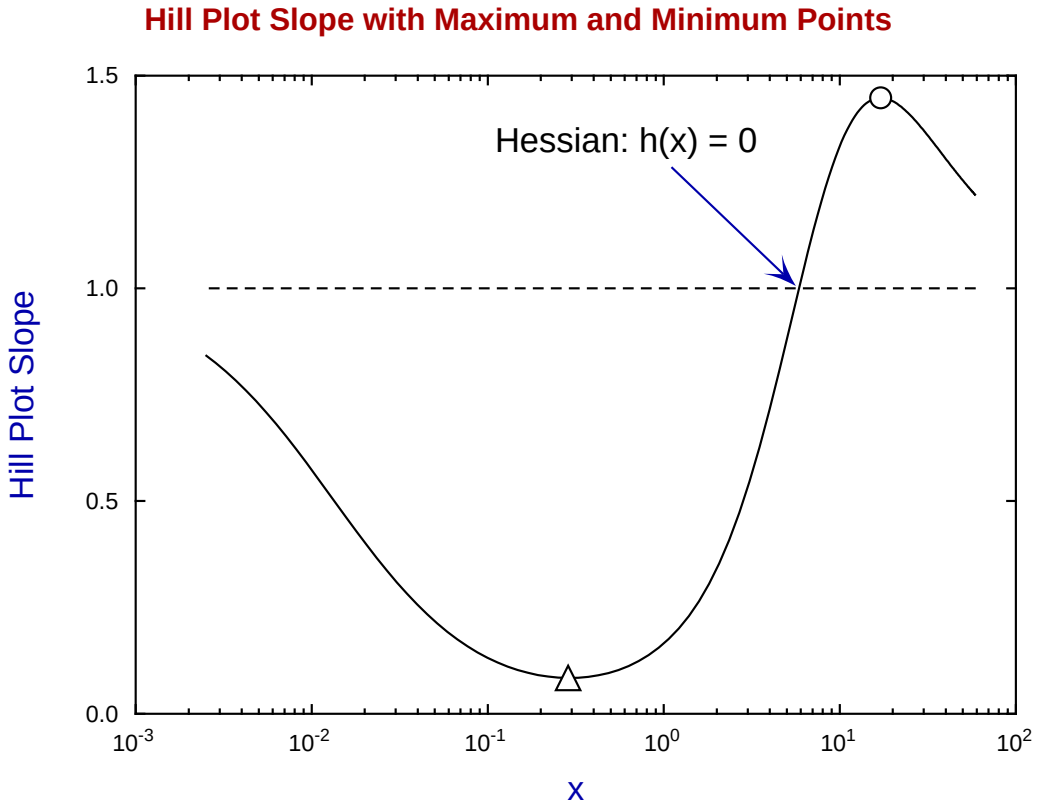
polynomial and Hessian are estimated, Hill slope is reported, and species fractions and binding isotherms are displayed, as is done automatically after every $n > 1$ fit by program **sffit**.

Cooperativity analysis

After fitting a model, program **sffit** outputs the binding constant estimates in all the conventions and, when $n > 2$ it also outputs the zeros of the best fit binding polynomial and those of the Hessian of the binding polynomial $h(x)$. The positive zeros of $h(x)$ indicate points where the theoretical one-site binding curve coinciding with the actual saturation curve at that x value has the same slope as the higher order saturation curve, which are therefore points of cooperativity change. The SimFIT cooperativity procedure allows users to input binding constant estimates retrospectively to calculate zeros of the binding polynomial and Hessian, and also to plot species population fractions.

For instance, for 4 sites with $K_1 = 100, K_2 = 10, K_3 = 1,$ and $K_4 = 0.1$, the Hessian has a positive zero at $x = 5.86139$, the minimum Hill slope in the range plotted is 0.0842, at $x = 0.28607$, the maximum is 1.44479, at $x = 17.059$, and the slope at half saturation is 1.0847, at $x = 6.5808$.

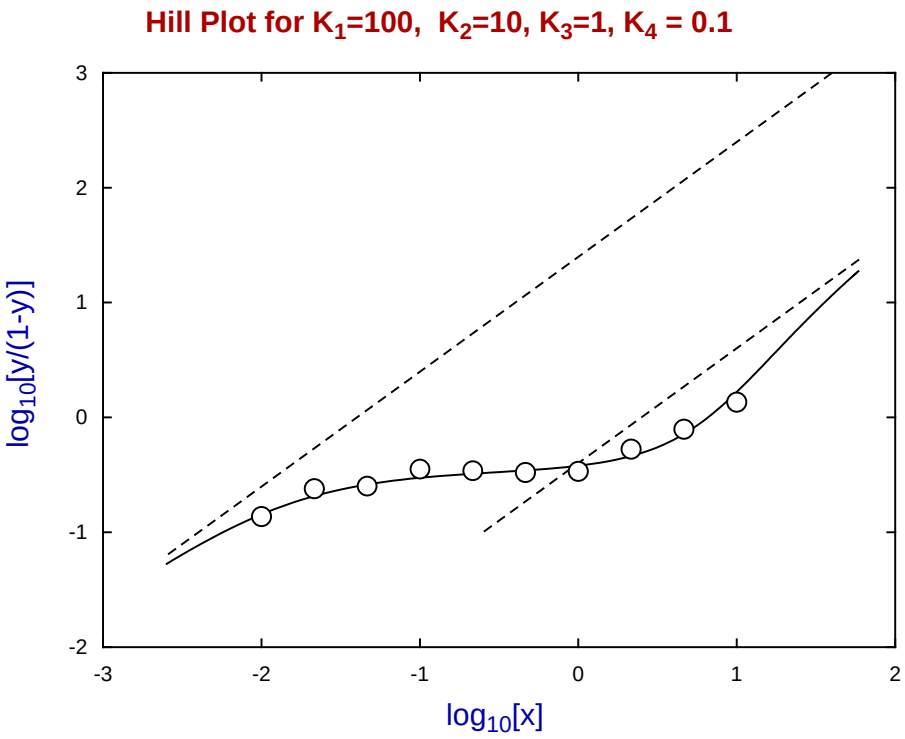
The next graph shows how the Hill plot slope varies with the maximum and minimum slopes indicated along with the point where the positive zero of the Hessian occurs.



The following graph shows the sort of complicated Hill plots that can be obtained when there are more than two cooperatively linked sites. The asymptotes are for the equation

$$y = \frac{kx}{1 + kx}$$

with $k = K_1/n$ as $x \to 0$ and $k = nK_n/K_{n-1}$ as $x \to \infty$.



References

[1] The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves.
Hill,A.V. (1910), *J.Physiol.* **40**, 4-7.

[2] The hemoglobin system. VI. The oxygen dissociation curve of hemoglobin.
Adair,G.S. (1925) *J.Biol. Chem.* **63**, 529-545.

[3] Allosteric Linkage.
Wyman,J. (1967), *J.Amer.Chem.Soc. 89.9* 2202-2218.

[4] Concerning the thermodynamic definition and graphical manifestations of positive and negative co-operativity.
Bardsley,W.G. & Wyman,J. (1978) *J. theor. Biol.* **72**, 373-376

[5] Critical points and sigmoidicity of positive rational functions.
Wood,R.M.W. & Bardsley,W.G. (1985) *Amer. Math. Month.* **92(1)**, 37-48

[6] Relationships between the magnitude of Hill plot slopes, apparent binding constants and factorability of binding polynomials and their Hessians.
Bardsley,W.G., Woolfson,R. & Mazat,J,-P. (1980) *J. theor. Biol.* **85**, 247-284

[7] Factorability of the Hessian of the binding polynomial. The central issue concerning statistical ratios between binding constants, Hill plot slope and positive and negative co-operativity.
Bardsley,W.G. & Waight,R.D. (1978) *J. theor. Biol.* **72**, 321-372

[8] Contributions of individual molecular species to the Hill coefficient for ligand binding by an oligomeric protein.
Edelstein,S.J & Bardsley,W.G. *J. Mol. Biol* (1997) **267**, 10–16