

Analysis of Hormone Binding Data, in "Methods In Diabetes Research" (1984),  
**Johnson, M.L.** and Frasier, S.G., Vol. I, (Larner, J. and Pohl, S., Eds.), Wiley and Sons,  
New York, pp. 45-61. Copyright ©1984 Wiley and Sons, Inc. This material used by  
permission of Wiley-Liss, Inc. a subsidiary of John Wiley & Sons, inc.

Visit John Wiley & Sons on the web at <http://www.wiley.com/>.

Buy a copy of this book online at <http://www.amazon.com>

D

# ANALYSIS OF HORMONE BINDING DATA

MICHAEL L. JOHNSON AND SUSAN G. FRASIER

*Diabetes Research and Training Center and Department of Pharmacology,  
University of Virginia School of Medicine, Charlottesville, Virginia*

1. Introduction	46
2. Graphical Methods and the Effects of Random Errors	47
3. Assumptions of Least-Squares Parameter Estimation	52
4. Evaluation of Confidence Intervals	54
5. Formulation of the Mechanism of Binding	56
6. Conclusion	60

---

## 1. INTRODUCTION

The primary objective of analyzing any type of hormone binding data is to determine a set of equilibrium constants and binding capacities that describe the process at equilibrium. Of equal importance, but overlooked by most investigators, is the accurate determination of the confidence interval, or standard error of the mean, for these parameters. Alterations of these determined parameters and their associated confidence intervals can then be used as a probe for the effects of various agonists and antagonists on the system.

These objectives are the same as those for the analysis of any type of ligand binding data. The graphical and mathematical methods employed for the analysis of hormone receptor interactions are also the same as those used for other ligand binding problems. Consequently, the analysis of hormone binding data is a subset of the larger problem of how to interpret the data for the interactions of molecules in solution.

In order to accomplish these objectives, three types of knowledge are required. First, and most important, is a knowledge of the types, distributions, and magnitudes of the random and systematic experimental errors inherent in the data. Second is the knowledge of how these experimental errors will affect the graphical or mathematical procedure used to evaluate the various parameters. Finally, the third requirement is a knowledge of the types of molecular interactions taking place during the binding process. The combination of the first two requirements will, in general, yield the optimal procedure for the analysis of the data. The third is required because all methods of parameter estimation make an assumption that the experimental data can be represented by some equation, and within this assumption, return the various parameters of that equation. This equation is a mathematical description of the molecular interactions.

If the estimates of the various experimental errors or the assumption of the molecular reaction mechanism are wrong, then the determined parameters and their associated confidence intervals will not correspond to what the investigator purports. Consequently, in a sense, all reports of ligand binding data should be required to have a section which begins, "These are the opinions upon which I shall base my facts" (1). This does not mean that the determined parameters (apparent affinity constants and binding capacities) are completely meaningless. An investigator can legitimately use these parameters, under defined conditions, as an experimental probe in the same way that a biologist might use muscle contraction, uterine weight, or behavioral changes as a measure of the effect of an agonist or antagonist on a system (2). It simply means that the investigator should clearly state all of the assumptions, conditions, and methods employed.

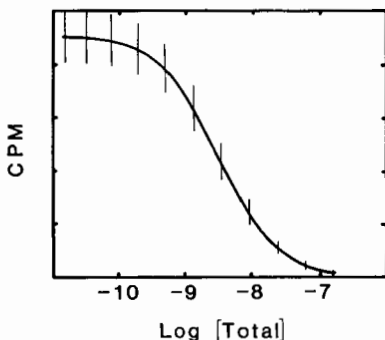
This chapter will discuss four major topics of particular relevance to the analysis of hormone receptor experimental data: (i) graphical methods and the effects of random experimental errors; (ii) the philosophy and assumptions of nonlinear least-squares techniques; (iii) the statistical methods for the evaluation

of confidence intervals; and (iv) the proper choice of the molecular mechanism for the hormone binding experiment. These sections will be discussed with particular reference to the effects of random and systematic experimental errors which might arise in the data. In each section specific examples will be employed to demonstrate the pertinent points.

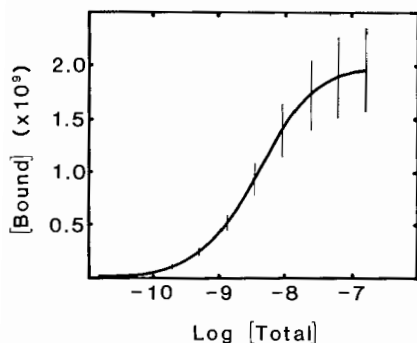
## 2. GRAPHICAL METHODS AND THE EFFECTS OF RANDOM ERRORS

As an example for the presentation of the various graphical methods, we will use a set of synthetic (or calculated) data. These data can then be perturbed by known amounts of "random" error varying with the total number of radioactivity counts, which demonstrates how this type of error affects the graphical procedure.

In a typical insulin binding experiment, a series of receptor preparations are incubated in the presence of a small amount of  $^{125}\text{I}$ -insulin plus varying concentrations of unlabeled insulin. After incubation, the preparations are fractionated by filtration or sedimentation, washed, and the amounts of radioactivity are measured by counting the filters or pellets, respectively. Figure 1 shows an example of such a synthetic experiment plotted as radioactivity bound versus the logarithm of the total molar concentration of insulin. The numerical values of the various constants are approximately those reported



**FIGURE 1.** Synthetic example of a typical insulin binding experiment (3). Data are simulated assuming two independent noninteracting classes of binding sites with  $K_a$ 's of  $2.0 \times 10^9 M^{-1}$  and  $2.1 \times 10^8 M^{-1}$ , and respective binding capacities of  $5 \times 10^{-10} M$  and  $1.5 \times 10^{-9} M$ . It should be noted that in the hormone receptor field, the binding capacities are usually expressed as per some amount of total protein or tissue preparation. For simplicity of simulation, we have expressed all concentrations in molarity. The error bars correspond to a radioactive counting error of 10 percent, when the total concentration is approximately  $10^{-11} M$  and increases to 20 percent as the number of counts decreases at higher unlabeled insulin concentrations. It is assumed that nonspecific binding did not occur in this data. It was further assumed that the concentration of added unlabeled insulin was known to have significantly better precision than the counting error.

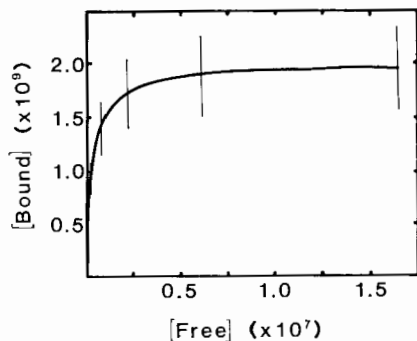


**FIGURE 2.** The data used in Figure 1 are corrected for the effective specific activity of labeled insulin and presented as moles bound versus the log of total insulin concentration in moles. The error bars in this figure, as well as Figures 3 to 5, are evaluated by simulating a data point at each end of the error bars in Figure 1 and then performing the appropriate transformations.

for the binding of insulin to liver plasma membranes (3) and are presented in the legend for Figure 1.

Data of this type are usually transformed from radioactivity bound into concentration bound, as is shown in Figure 2. This transformation is accomplished by dividing the bound counts of radioactivity by the effective specific activity of the combination of labeled and unlabeled insulin. It is interesting to observe the way that the radioactive counting error is propagated from Figure 1 to Figure 2. This increasing experimental error at higher concentrations of unlabeled hormone arises from the combination of two phenomena. First, at higher concentrations of unlabeled hormone, the number of radioactive counts is quite low. Under these conditions most investigators do not count to a constant number of counts above background, that is, a constant counting error. Most investigators will let the counting error increase in order to decrease the amount of counting time required. Second, at high concentrations of cold ligand, the effective specific activity approaches zero. Since the bound concentration is calculated by dividing this poorly determined low number of counts by a specific activity which approaches zero, the absolute error in the bound concentration increases markedly.

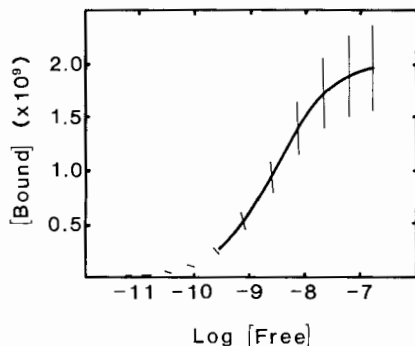
It should be noted that most of the mathematical formulations for ligand binding problems are presented in terms of the free (or unbound) concentration of ligand (or hormone in this case), rather than the total concentration. Accordingly, the free concentration is then calculated as the total concentration less the bound concentration. Typically, the data are graphed showing bound versus either the free or the logarithm of the free concentration, as in Figures 3 and 4. Presenting the data as bound versus free, as in Figure 3, places most of the data points on the first few percent of the left side of the graph, and therefore is not particularly useful for visual interpretation. Figure 3 is, however, the formulation usually employed for ligand binding problems aside from the hormone receptor field. The logarithmic representation of Figure 4 is more useful for visual inspection. It is quite interesting that when the concentration of free hormone falls below  $10^{-8} M$  in this particular example (Figure 4), the error bars are no longer vertical. This happens as a result of the method used to calculate the free concentration from the bound concentration. At these low



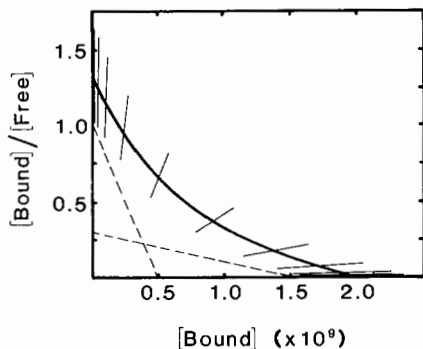
**FIGURE 3.** The data of Figure 1 are presented as amount bound versus free molar concentrations. See Figures 1 and 2 for details.

total concentrations the bound concentration can, in some cases, be a significant fraction of the total concentration. When this occurs, as in this example, any small experimental error in the bound concentration will generate an error in the free concentration, since the free concentration is calculated as total concentration minus the bound concentration. Thus, since both the bound and free concentrations have the same associated errors, the error bars will no longer be vertical. The same nonvertical error bars occur in Figure 3 but are less obvious due to the format of the data presentation.

Probably the most commonly used and abused graphical procedure for the analysis of hormone binding data is the Scatchard plot. If the chemical mechanism of binding the hormone to the receptor can be described as a single class of binding sites, then the Scatchard plot will be a straight line. The same simulated data set is again presented in Figure 5 as the bound divided by the free concentrations versus the bound concentration (the Scatchard transformation). Also shown in Figure 5 are the Scatchard plots for the two individual classes of binding sites, represented by the dashed lines. It is interesting to note that the error bars in this formation seem to radiate from the origin. On the right side of this graph the total concentration of unlabeled hormone is very high and well determined, while the bound concentration is significantly



**FIGURE 4.** The data of Figure 1 are presented as amount bound versus the log of free molar concentration. The calculated line has been omitted at low concentrations so that the angle of the error bars is easily observable. See Figures 1 and 2 for details.



**FIGURE 5.** Scatchard plot (bound divided by free molar concentrations versus bound molar concentration) for the data in Figure 1. The dashed lines correspond to the Scatchard plots of the individual classes of binding sites. See Figures 1 and 2 for details.

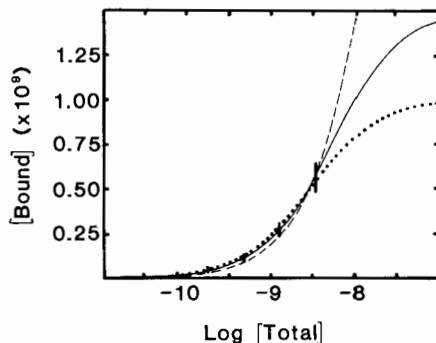
lower and poorly determined. Under these conditions, the bound-to-free concentration ratio ( $Y$  axis) will be very small with a small absolute error, while the bound concentration ( $X$  axis) will be large, and the error bars will be almost horizontal. At extremely low total concentrations (the left side of this graph), the free concentration will be very small. Thus, the bound-to-free concentration ratio will be very large and have a large absolute error compared to the bound alone, and the error bars will appear to be almost vertical.

There is currently an ongoing discussion in the literature (2, 4, 5) as to which of these plots (Figs. 1–5) is the proper method to analyze and present hormone binding data. Munson and Rodbard (2) have pointed out that “the statistical information content of the data is not altered by presentation in one or the other coordinate system: a simple algebraic manipulation will convert one format into the other.” This is obviously true. However, as pointed out by Klotz (4, 5), what is altered is our perception of the data. Klotz presented several examples from the literature (4, 5) in which various investigators have not collected data at high enough concentrations to reach the inflection point of the plot of bound versus the logarithm of the free concentration. These investigators presented their data as a Scatchard plot without noting that the amount bound had not reached saturation. Klotz points out that under these conditions it is impossible to determine either the binding capacity or the equilibrium constant from the data. Figure 6 is an example of the point which Klotz was attempting to make. In this figure data were simulated for a single class of binding sites with a total concentration range of  $10^{-11} M$  to  $3 \times 10^{-9} M$ . Since the  $Kd$  for this simulated data was  $2 \times 10^{-8} M$ , with a corresponding binding capacity of  $1.5 \times 10^{-9} M$ , these simulated data only reach approximately 30 percent of saturation. The solid line in Figure 6 corresponds to the correct answers. The dotted and dashed lines correspond to binding capacities of  $1.0 \times 10^{-9} M$  and  $5 \times 10^{-9} M$ . Note that all of these appear to describe the data over their limited range. This is a common problem when the Scatchard plot is used. The best approach is to use several different graphical representations.

Neither Klotz nor Munson and Rodbard noted what is perhaps the most serious abuse of the Scatchard plot. If the Scatchard plot is not a straight line,

then either the underlying molecular mechanism is not a single class of non-interacting binding sites or the nonspecific binding has been incorrectly treated. In the case of nonlinear Scatchard plots, some investigators are tempted to draw two straight lines through the limiting slopes of the Scatchard plot and assume that these slopes reflect the binding affinities of the high- and low-affinity classes of sites. First, these investigators are assuming that two classes of sites exist, although the evidence may be insufficient to support this idea. Second, it is nearly impossible to perform such an operation with reasonable precision or statistical validity. Third, and perhaps most importantly, when the two classes of site have affinities which are reasonably close, such as within a single order of magnitude, as shown in our simulated Scatchard plot (Fig. 5), the limiting slopes do not correspond to the individual binding affinities. This is obvious in Figure 5 by a comparison of the dashed lines and the limiting slopes.

Munson and Rodbard (2) have pointed out that "One may avoid the issue of graphical methods altogether, by the use of computerized nonlinear least-squares curve fitting." This statement, while true, could also be misleading. The formulation of any problem for nonlinear least-squares analysis assumes that some observable parameter, possibly a bound concentration, is a function of one or more other observable parameters, such as the free concentration. Consequently, irrespective of whether the graph is actually drawn, a "virtual graph" is constructed within the computer program used for the analysis. This virtual graph is subject to many of the same problems as the more classic graphical methods. We noted earlier that the various graphical presentations alter our perception of the data. In like manner some of the virtual graphical representations of the computer-analyzed data violate the basic assumptions of the method of least-squares and are thus not allowable. Hence, to be able



**FIGURE 6.** A simulated experiment showing the consequences of failing to approach a saturation level. Data were simulated as in Figure 1 except we assumed a single class of binding sites with  $Kd$  of  $5 \times 10^{-9} M$  and a binding capacity of  $1.5 \times 10^{-9} M$ . The solid line corresponds to the correct answers, the dashed line corresponds to a  $Kd$  of  $2.15 \times 10^{-8} M$  and a binding capacity of  $5 \times 10^{-9} M$ , and the dotted line corresponds to a  $Kd$  of  $2.47 \times 10^{-9} M$  and a binding capacity of  $1 \times 10^{-9} M$ .

to apply this technique, it is important to be familiar with the basic assumptions of the least-squares method.

### 3. ASSUMPTIONS OF LEAST-SQUARES PARAMETER ESTIMATION

It is probably not critical for a biochemist to be familiar with all of the detailed mathematics of the nonlinear least-squares methods of parameter estimation. These numerical procedures have been published elsewhere (6–10). The important point is a knowledge of the assumptions and underlying philosophy of the method.

The basic method for parameter estimation takes a function  $G$  and a series of data points,  $X_i$  and  $Y_i$ , and determines a set of parameters, such as binding capacities and equilibrium constants. These parameters, denoted by the vector  $\alpha$ , are determined such that the sum of the squares of the residuals, that is, the difference between the function and the data points, is a minimum. The function  $G$  is the mathematical representation for the virtual graph mentioned earlier. Examples of this function will be discussed later in this review. In practice, this procedure is simply an algorithm which, when given an initial guess for the parameters,  $\alpha$ , will find a better estimate for those parameters. The procedure is then applied in an iterative fashion until the parameters being determined do not change within some specified tolerance.

Linear least-squares is a special case of the more general nonlinear least-squares method. Linear does not imply a straight line in this context since higher-order polynomials can also satisfy the definition. A system is referred to as linear when all of the second derivatives of a function with respect to the parameters being estimated are zero. For example, if  $G$  represents a quadratic equation of the form  $a + bx + cx^2$ , then the second derivatives with respect to  $a$ ,  $b$ , or  $c$  will all be zero. The important item to note is that the only difference between linear and nonlinear least-squares techniques is the number of iterations required to perform the analysis because the assumptions and problems associated with nonlinear least-squares parameter estimation are also present in linear least-squares analysis.

The least-squares parameter estimation method makes the following assumptions:

1. All of the experimental error of the data can be attributed to the ordinate (the  $Y$  axis).
2. The random experimental errors of the data can be described by a Gaussian (bell-shaped) distribution.
3. No systematic error exists in the data.

All three of these assumptions must be satisfied for the least-squares procedure to be valid. Furthermore, each of the assumptions has marked implications as to the “proper” method to proceed in analyzing hormone binding data.

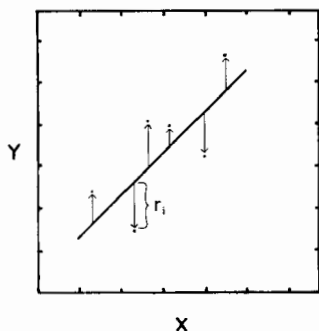


FIGURE 7. Graphical depiction of the method of least-squares parameter estimation.

The first assumption is graphically depicted in Figure 7. The least-squares method picks values for the various parameters of the function such that the sum of the squares of the vertical distances,  $r_i$ , is a minimum. It should be noted that it is only the vertical distances that are minimized, not the horizontal or perpendicular distances. Therefore, in order to be able to correctly use a least-squares analysis, the problem must be formulated in such a manner that the precision of the determination of the values along the abscissa ( $X$  axis) must be significantly greater than the precision of the ordinate ( $Y$  axis). In general, there is no method to circumvent this requirement. It cannot be done by "appropriate weighting factors" for a nonlinear problem.

The second assumption means that if a given data point on Figure 7 was experimentally measured an infinite number of times, the resulting distribution of values (or  $r_i$ 's) could be described by a Gaussian distribution. This is a reasonable assumption for a large number of, but not all, experimental procedures. New statistical techniques are currently being developed to circumvent this assumption, but their application to nonlinear problems has as yet not been fully investigated (cf. ref. 11). The second assumption also implies that the number of data points is sufficient to insure a good statistical sampling of the random error. The exact number of data points required to meet this assumption is difficult to specify. However, what can be stated is that there is no practical difference between experimentally determining 6 data points distributed within a given range of concentration three times or measuring 18 separate data points within that range. Furthermore, if 6 data points were measured in triplicate, the analysis should be performed on the 18 individual values, not 6 average values.

The third assumption clearly states that systematic errors cannot exist in the data. The only way to circumvent this requirement is to include a series of terms in the function  $G$  to describe the systematic errors. For example, nonspecific binding is a source of systematic error for the determination of specific binding. It was previously been shown (9) that the function  $G$  can be written in such a manner that nonspecific binding is explicitly included and is thus no longer a systematic error of the data. An example of this procedure will be given later in this review.

As previously described, all computer programs that perform a least-squares analysis of experimental data generate a virtual graph of the data internal to

the program. The exact choice of graphical formulation is critical to the validity of the least-squares method employed because some of them violate the basic assumptions of the method.

The first assumption listed for the least-squares method was that all of the errors of the data can be attributed to the ordinate ( $Y$  axis). An examination of Figures 1–5 indicates that this assumption is violated by the Scatchard plot (Fig. 5) and by the plots of bound versus the free (Fig. 3) and bound versus the logarithm of the free concentration (Fig. 4). This is a consequence of the method of calculating the free concentration from the total and bound concentrations. It thus appears that the only valid formulations are plots of amount bound or radioactivity versus the total concentration or the logarithm of the total concentration. This is the formulation used by the computer program of Munson and Rodbard (9).

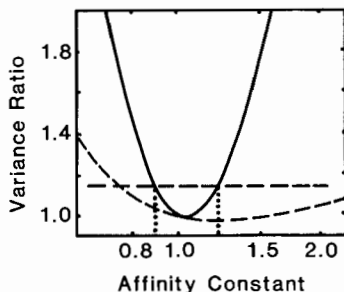
At first it might seem that this problem could be circumvented by a direct measurement of the free concentration. However, the usefulness of this approach is doubtful because the free concentration will generally correspond to a very low level of radioactivity and thus will be subject to large errors in measurement. In this case it is necessary for the investigator to allow the number of counts at the low end of the cpm scale to remain high, even though this requires a large amount of counting time.

The data of Figures 1–5 were simulated assuming an increased radioactive counting error when proportionately greater amounts of unlabeled hormone are added. This is a standard approach to such experiments. It is important to be aware that the predicted distribution of experimental error due to counting uncertainties is not Gaussian; it is a Poisson distribution. At large numbers of counts, the Poisson distribution can be approximated by a Gaussian curve, but not as the number of counts approaches zero. This means that if an investigator wants to apply least-squares with some measure of validity at low specific activities, that is, high concentrations of unlabeled hormone, it is essential that the investigator count all samples to a predetermined number of counts (e.g., 3000) rather than to a predetermined time.

#### 4. EVALUATION OF CONFIDENCE INTERVALS

Munson and Rodbard (2, 9) have stated that nonlinear least-squares provides a measure of the standard errors and confidence limits for each of the estimated parameters. This is true only if the investigator is willing to make some assumptions. There is no exact analytical method to evaluate the confidence intervals from a nonlinear least-squares analysis (10).

Most of the generally available computer programs employ a method for the evaluation of confidence intervals based on the method used for linear least-squares fitting (12). This method assumes that the parameters are not



**FIGURE 8.** A simple single parameter least-squares analysis of a rectangular hyperbola, showing the confidence region as evaluated using the variance ratio. Any value of  $K$  predicting a variance ratio less than a critical value, which is determined by the number of degrees of freedom and confidence probability, will lie within the confidence region. On this graph, for example, the intersection of the variance ratio at 1:14, denoted by the horizontal dashed line, and the data set indicated by the solid curve, projects the confidence interval onto the  $X$  axis. See text for discussion of data.

correlated with each other, that the variance space\* can be predicted from its curvature at the point of minimum variance, and that the answer can be approximated by a linear least-squares fit. The consequences of these assumptions would indicate that the computer program will predict a symmetrical confidence interval, which is an underestimate of a true confidence interval. It is relatively easy to demonstrate that the true confidence interval for a nonlinear problem is not symmetrical (7, 10) and thus cannot be described as plus or minus a standard error.

It has been previously shown that the true confidence intervals of the fitted parameters are defined by the  $F$  statistic calculated from the ratio of the variance at any point in space to the variance at the minimum of that space (6, 13). Any particular set of parameters which predicts an  $F$  statistic below a particular critical value is within the acceptable confidence intervals for those parameters. The critical value of this  $F$  statistic is determined by the desired confidence probability, the number of data points, and the number of parameters being estimated. This is true for any functional form, no matter how nonlinear and asymmetrical the variance space.

An example of the use of this procedure to find the confidence interval for a one-parameter least-squares estimation is given in Figure 8. This example is a simulation of a simple ligand binding problem where the number of binding sites is known independently and there is no cooperative interaction of any type between the binding sites. With these assumptions the fractional occupancy can be expressed as  $K[X]/(1 + K[X])$  where  $K$  is the equilibrium binding

\*In this context the variance space is defined as a multidimensional description of the value of the variance for any values of the parameters being determined. The curvature of the variance space is determined by the second derivative of the variance with respect to each of the parameters being determined.

constant of the ligand and  $[X]$  is the free concentration of the ligand. This is the commonly used functional form for the evaluation of hormone receptor interactions.

Each curve of Figure 8 was calculated by generating 50 data points with equal logarithmic spacing in  $[X]$ . The data were then perturbed with Gaussian distributed pseudorandom error with standard deviation of 0.05. The variance, which corresponds to any value of  $K$ , can be calculated from the residuals, that is, the difference between the data points and the function evaluated at the corresponding free ligand concentration. The variance ratio,  $F$  statistic, corresponding to a given value of  $K$  can then be calculated as the ratio of the variance for that particular value of  $K$  and the variance evaluated at the best estimated value for  $K$ . The solid line in Figure 8 was generated with data where the free ligand concentration varied over four orders of magnitude centered on the simulated dissociation constant:  $0.01 < [X] < 100$ . The critical value of the  $F$  statistic for a 67 percent confidence interval is 1.14 for this example. This predicts that the equilibrium constant from these data is 1.03 with a one standard deviation confidence interval of 0.89–1.21. However, it should be noted that in real experiments, data are rarely obtainable over a perfectly selected four orders of magnitude in concentration. Consequently, the dashed curve in Figure 8 was calculated by the same procedure, with the range of  $[X]$  being only one order of magnitude in concentration and not centered around the dissociation constant,  $10 < [X] < 100$ . The same critical value of the  $F$  statistic predicts that the equilibrium constant is 1.20 with a one standard deviation confidence interval of 0.75–2.79. It is interesting to note that this pseudorealistic one-parameter curve-fitting problem can yield very asymmetrical confidence intervals. Such asymmetrical confidence intervals can only be evaluated by searching the variance space for the critical value of the  $F$  statistic (10). The classical methods of evaluating confidence intervals assume that the dashed line in Figure 8 can be approximated as a symmetrical parabola centered about the best approximation of the answer. Consequently, they will predict that the confidence interval corresponding to the dashed line in Figure 8 is symmetrical.

More complex methods for the evaluation of confidence intervals are available (6–8, 10, 14). These methods require a significant increase in computer time, but as is shown in Figure 8, the results are worth the expenditure. The investigator should be cautious of any computer program that expresses the confidence interval as plus or minus a single standard error, since in practice the confidence interval is asymmetrical.

## 5. FORMULATION OF THE MECHANISM OF BINDING

The only remaining assumption for the utilization of nonlinear least-squares techniques is a knowledge of the particular molecular mechanism of hormone binding. This assumption is required because the nonlinear least-squares tech-

nique requires an exact explicit function  $G$ , which relates the dependent variables with the independent variables: that is, the bound concentration versus the logarithm of the total concentration, for our particular case.

Typically, the assumption is made that the molecular mechanism of interest is an unknown number of independent noninteracting binding sites. With this assumption the correct functional form to describe the bound concentration as a function of the free concentration will be

$$G(\text{free}) = [R] \sum_{i=1}^m \frac{n_i [\text{free}]}{Kd_i + [\text{free}]} \quad (1)$$

where  $G(\text{free})$  describes the bound concentration of hormone,  $R$  is the concentration of receptor molecules,  $i$  is an index which refers to each of the  $m$  different classes of binding sites,  $n_i$  is the number of binding sites in each of the  $i$  classes of sites, and  $Kd_i$  is the corresponding dissociation constant for each of the classes of binding sites. As shown earlier, a virtual graph of bound versus free concentration violates the basic assumptions of the least-squares method of parameter estimation. In order to properly utilize the least-squares technique, the data must be represented as a function of the total concentration. This can be accomplished by noting that the total concentration is a function of the free, that is,

$$[\text{total}] = [\text{free}] + G(\text{free}) \quad (2)$$

For any particular set of values of  $m$ 's,  $n_i$ 's,  $kd_i$ 's, and total concentrations, Equation (2) can be numerically solved to find the free concentration of hormone. This free concentration can then be used to evaluate the bound concentration,  $G(\text{free})$ , at any value of the total concentration. There are a few available computer programs that use this strategy (9, 10).

The most common definition for nonspecific binding is the unsaturable low-affinity specific binding which the investigator is not currently interested in studying. The key words in this definition are *unsaturable* and *specific binding*. By acknowledging that the nonspecific binding is actually a low-affinity specific binding, it can be incorporated into Equation (1) by simply adding an additional class of binding sites. A word of caution is in order regarding the key term, *unsaturable*. Unsaturable implies that the  $Kd_i$  for the nonspecific binding is always significantly greater than the free concentration. When this occurs, the corresponding term for nonspecific binding in Equation (1) degenerates to the form

$$\frac{n_i}{Kd_i} [\text{free}] \quad (3)$$

It is important to note that this form of the binding equation implies that  $n_i$  and  $Kd_i$  are always correlated with each other. Unique values of  $n_i$  and  $Kd_i$  can not be found, but their ratio is determinable and can be substituted into Equation (1) to yield the form

$$G(\text{free}) = K_{\text{ns}} [\text{free}] + [R] \sum_{i=1}^m \frac{n_i [\text{free}]}{Kd_i + [\text{free}]} \quad (4)$$

where  $K_{\text{ns}}$  is an arbitrary constant which describes the nonspecific binding: that is,  $K_{\text{ns}} = n_i/Kd_i$ . This is basically the method employed by Munson and Rodbard (9).

To recapitulate, the least-squares method for computer analysis generates a "virtual" graph of the bound concentration versus the logarithm of the free concentration. This graph is then analyzed by utilizing a combination of Equations (1)–(4). By this method, the nonspecific binding is determined simultaneously with the specific binding.

At this point we will present a few examples of the procedure. These examples will be chosen to address the question "What happens if the assumed molecular mechanism is incorrect?" One example is shown in Table 1 and Figure 9. The data are the same simulated data as were presented in Figures 1–5. These data correspond to two classes of binding sites with no nonspecific binding. These two classes differ by an order of magnitude in affinity with the lower affinity class having three times the binding capacity. The exact constants are presented in Table 1 as model 1. The dashed line in Figure 9 (model 2 in Table 1) corresponds to an analysis of these data assuming a single class of binding sites. The dotted line in Figure 9 (model 3 in Table 1) corresponds to the analysis of these data as a single class of binding sites and 0.32 percent nonspecific binding. The solid line in Figure 9 (model 4) assumes two classes of binding site with an equal binding capacity. All of these models lie within the expected error bars of the data points. An inspection of Table 1 shows that the affinity constants determined for models 2–4 are quite different from those of model 1. A visual inspection of Figure 9 would indicate that the only way to distinguish these four models would be to collect data at even higher concentrations and with greater precision.

Several groups of workers have suggested that receptor aggregation, clustering, or binding of a second protein are important for the functioning of the

**TABLE 1. Values of Equilibrium Constants Employed for the Calculated Lines in Figure 9**

Model	$N_1$	$K_1^a$	$N_2$	$K_2^a$	$K_{\text{ns}}^a$
1 <sup>b</sup>	$1.50 \times 10^{-9}$	$2.10 \times 10^8$	$0.50 \times 10^{-9}$	$2.0 \times 10^9$	0
2 <sup>c</sup>	$1.73 \times 10^{-9}$	$6.79 \times 10^8$	0	0	0
3 <sup>d</sup>	$1.55 \times 10^{-9}$	$7.75 \times 10^8$	0	0	0.0032
4 <sup>e</sup>	$1.02 \times 10^{-9}$	$0.95 \times 10^8$	$1.02 \times 10^{-9}$	$1.14 \times 10^9$	0

<sup>a</sup>  $K$ 's are presented as association constants.

<sup>b</sup> Model 1 was used to simulate the data points for Figures 2 and 9, as described in the legend to Figure 1.

<sup>c</sup> Model 2 corresponds to the dashed line in Figure 9.

<sup>d</sup> The dotted line in Figure 9 was calculated from model 3.

<sup>e</sup> The solid line is model 4.

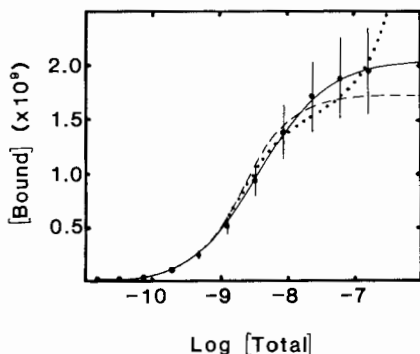


FIGURE 9. The data from Figure 2 analyzed by three different models. The actual values of the constants are given in Table 1.

insulin receptor (14–16). If such a phenomenon is present, it should be included in the mathematical formulation of the binding, that is, Equations (1)–(4). The Scatchard plot for the binding of a ligand to a binding protein, which is simultaneously undergoing a self association, is shown in Figure 10. The particular example was calculated to simulate the binding of rhodamine 6G (the ligand) to glucagon (the binding protein). At pH 10.6 in 0.6 *M* phosphate, glucagon exists at an equilibrium between a monomer and a trimer with an equilibrium constant of  $4.8 \times 10^7 M^{-1}$  (17). Rhodamine 6G binds to the trimer with an affinity of  $9.4 \times 10^4 M^{-1}$  but does not bind to the monomeric form (17). Glucagon is thus a model for the binding of hormones, in this case rhodamine 6G, to a receptor which polymerizes, in this case glucagon. It should be noted that the shape of the Scatchard plot in Figure 10 would seem to indicate that multiple classes of binding sites, some form of heterogeneity, or negative cooperativity is taking place. But remember also that these data were simulated assuming a single binding site with no heterogeneity or cooperativity other than the concomitant self-association of the receptor.

Figure 11 presents an analysis of the same data as in Figure 10, assuming that the molecular mechanism is two classes of binding site. The apparent binding affinities from this analysis are  $1.79 \times 10^3 M^{-1}$  and  $40.7 M^{-1}$  with

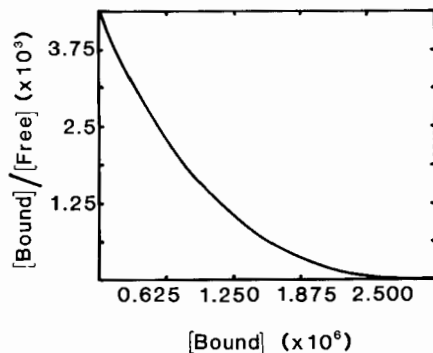
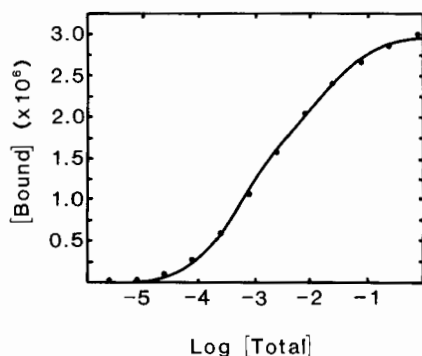


FIGURE 10. Simulated Scatchard plot for the binding of rhodamine 6G to 10  $\mu M$  glucagon at pH 10.6, 0.6 *M* phosphate (17).



**FIGURE 11.** The simulated data for rhodamine 6G binding to glucagon (as in Fig. 10) presented as the bound versus the logarithm of the total concentration. The solid line was calculated for a model corresponding to two classes of binding sites.

corresponding binding capacities of 1.8 and 1.2  $\mu M$ . It is apparent that these constants are significantly different from the actual values. However, observe that the calculated solid line in Figure 11 is easily capable of describing the data.

Both of these latter cases were specifically chosen to demonstrate a single point: The analysis of ligand binding data can yield multiple sets of answers depending upon what assumptions are made for the analysis.

## 6. CONCLUSION

The primary motivation for this review has been to acquaint the reader with some of the methods for the analysis of hormone binding data, as well as the assumptions, advantages, and disadvantages of those methods. The reader may, by this time, have the impression that it is impossible to attain a "correct" answer from the analysis of hormone binding data. This is only true if the investigator fails to state the assumptions and conditions under which the measurements and analysis were performed. The parameters determined by these methods will be correct within these assumptions.

The reader should note that up to this point no mention has been made of experimental problems such as degradation of the radioactive label, or of other processes such as hormone and/or receptor internalization. The reader is reminded that the methods of analysis assume an equilibrium process and, as a consequence, kinetic phenomena such as the degradation of label and internalization cannot be treated. The investigator could collect all of the binding data as a function of both concentration and time. This should, in principle, allow the investigator to write a series of kinetic equations (i.e., differential equations) for the function  $G$  and include processes such as degradation of label or receptor internalization. In practice, these mathematical procedures are very difficult to perform and require vast amounts of data as well as computer time. In short, this procedure is not for the uninitiated!

Some of the specific examples presented in this review show that the method of data analysis requires some modification of the commonly used protocols

for the collection of experimental data. Specifically, the investigator should be certain that the concentrations have reached saturation. In addition, the investigator should be particularly careful about counting error at high concentrations of unlabeled hormone; that is, the investigator should set the counter for a constant number of counts above background, not a constant time.

It may be, as noted by Klotz (5), that it is impossible to determine the true mechanism and parameters from hormone binding experiments alone. Munson and Rodbard (2) have pointed out that a large amount of information can be obtained without actually determining the true mechanism of hormone binding by simply considering the determined parameters as "apparent" values. These apparent values can still be used as a probe of the hormone receptor interaction without the assumption that they are actually correct.

## REFERENCES

1. J. Faulkner, *Nature* **253**, 231 (1975).
2. P. J. Munson and D. Rodbard, *Science* **220**, 979 (1983).
3. C. R. Kahn, P. Freychet, J. Roth, and D. M. Neville, Jr., *J. Biol. Chem.* **249**, 2249 (1974).
4. I. M. Klotz, *Nature* **217**, 1247 (1982).
5. I. M. Klotz, *Nature* **220**, 981 (1983).
6. G. E. P. Box, *Ann. N.Y. Acad. Sci.* **86**, 792 (1960).
7. M. L. Johnson, H. R. Halvorson, and G. K. Ackers, *Biochem.* **15**, 5363 (1976).
8. M. L. Johnson, J. J. Correia, D. A. Yphantis, and H. R. Halvorson, *Biophys. J.* **36**, 575.
9. P. J. Munson and D. Rodbard, *Anal. Biochem.* **107**, 220 (1980).
10. M. L. Johnson, *Biophys. J.* **44**, 101 (1983).
11. P. J. Huber, *Robust Statistics*, Wiley-Interscience, New York, 1981.
12. F. B. Hildebrand, *Introduction to Numerical Analysis*, McGraw-Hill, New York, 1956, p. 82.
13. L. Endrenyi and F. H. F. Kwong, *Acta Biol. Med. Germ.* **31**, 495 (1973).
14. C. R. Kahn and K. L. Baird, *Physical Chemical Aspects of Cell Surface Events in Cellular Regulation*, Elsevier/North-Holland, New York, 1979, p. 119.
15. C. DeLisi, *Physical Chemical Aspects of Cell Surface Events in Cellular Regulation*, Elsevier/North-Holland, New York, 1979, p. 261.
16. J. T. Harmon, C. R. Kahn, E. S. Kempner, and M. L. Johnson, *Current Views on Insulin Receptors*, Academic Press, New York, 1981, p. 37.
17. M. L. Johnson, S. Formisano, and H. Edelhoch, *J. Biol. Chem.* **253**, 1353 (1978).