

for *p*-xylene indicate no ionization potential lower than about 8.2 eV.³⁹ This same study³⁸ also obtained a much lower a_{2u} ionization potential for benzene than that calculated by Newton, Boer, and Lipscomb,²⁶ and on that basis questioned the calculations. We must concur with the criticism and hence have, for instance, assumed the transition energy of the $\omega_5 \rightarrow 2b_{1g}$ transition in *p*-xylene-TCNE (Table IX) to be 2.70 eV (this effectively pulls orbital ω_5 down to -8.19 eV, within the region expected from ionization potential studies). The errors in the calculation, however, are rather serious and allow us to make no comments on the effects of increased methylation in molecular complexes.

Three very recent works have been brought to our attention dealing with the problem studied here.^{36,37,40} Mantione has calculated the inductive contribution to the dipole moments³⁷ and the energy of formation³⁶ of a rather large number of TCNE-aromatic complexes. The form of the perturbation theory used by her does not contain a charge-transfer term and is probably equivalent to (25), although with a different form of the

(38) A. D. Baker, D. P. May, and D. W. Turner, *J. Chem. Soc., B*, 22 (1968).

(39) The original paper of Newton, Lipscomb, and Boer²⁶ calculated an ionization potential for *p*-xylene of 6.98 eV, better than the present calculation but still subject to the same criticism. The difference in calculations is due to a different choice of diagonal U_{ii} matrix elements.

(40) R. J. W. Le Févre, D. V. Radford, and P. J. Stiles, *J. Chem. Soc., B*, 1297 (1968).

repulsion. Her method of evaluating electrostatic terms is much like ours except that she does not include π -quadrupole effects. In fact, comparable terms for benzene-TCNE in the Mantione work³⁷ are almost identical with those calculated here. Neglect of π quadrupoles, however, gives $\Delta E_f = -4.80$ kcal/mole for benzene-TCNE, as compared to the present $\Delta E_f = -6.27$ and measured $\Delta E_f = -7.36$. Neglect of π quadrupoles would seem important.

Le Févre, Radford, and Stiles⁴⁰ have attempted to evaluate the amount of charge transfer in the ground state of molecular complexes by attributing that part of the measured dipole moment not accounted for by the induced moment as being due to charge transfer, and hence a measure of the amount of charge transfer in the ground state. Their conclusion, that van der Waals forces rather than charge-transfer interactions seem primarily responsible for the ground-state dipole moment, is not in disagreement with our results. We do feel, however, that the present state of experimental complex dipole moments does not warrant use of this procedure as an accurate measure of amount of charge transfer.

Acknowledgment. We wish to thank Dr. P. O'D. Offenhartz for making several computer subroutines available to us and for invaluable assistance during the debugging of the programs. We also wish to thank Dr. D. R. Williams for a number of helpful discussions.

Theory of the Measurement of Weak Molecular Complexes.

I. General Considerations

David A. Deranleau

Contribution from the Laboratorium für Molekularbiologie, Eidgenössisch Technische Hochschule, 8006 Zürich, Switzerland. Received September 30, 1968

Abstract: Some aspects of the measurement of formation constants and other constants for weak molecular complexes are considered from the standpoint of basic binding theory. The theoretical minimum errors in the formation constant k and the extinction coefficient ϵ are discussed as a function of the saturation fraction of the most dilute component, and it is shown in agreement with previous work that the most accurate values of either parameter are obtained when the saturation fraction lies between 0.2 and 0.8. Outside this region, the determined values become extremely uncertain, and in addition, insufficient data are available to adequately fit a given stoichiometric model. It is suggested that the criteria that the highest concentration of the excess component be not less than $0.1/k$ are insufficient to assign a given stoichiometric model to the data, but apply instead to the smallest range of saturation fraction required to show that a line of finite limiting slope and/or intercept exists. To develop criteria by which a given stoichiometric model can be considered in 1:1 correspondence with the phenomenological equation fitting the data, the principles of information theory are applied to the binding process. It is concluded that of the order of 75% of the saturation curve is required to show the correspondence between the equation of the model and the equation fitting the data. Also discussed are various plotting forms and the errors arising from various rearrangements of the basic binding equation.

The problem of the determination of formation constants and other constants for weak intermolecular complexes has received considerable attention during the past several years, and for good reason, since the proof of existence of such complexes depends largely on the obtention of unique values for the formation constant k and the extinction coefficient ϵ . The most critical discussion is that given by Person,¹ who points

out what some authors seem to have missed: that the most accurate values of the formation constant are obtained when the concentration of the complex is approximately the same as the free concentration of the most dilute component. Although Person limited his discussion to a consideration of the errors arising when the concentrations fell below or above certain limits,

(1) W. B. Person, *J. Am. Chem. Soc.*, **87**, 167 (1965).

resulting in the inability to determine either k or ϵ separately,² the central idea has more far reaching consequences in terms of the assignment of an adequate model for the observed interaction. In order for limit criteria such as presented by Person to be applicable, it is of primary concern to determine that the model from which the limit criteria were derived is in fact the model correctly describing the experimental situation. A further complication arises when the limit criteria depend on a prior knowledge of k , a quantity determined in the experiment and, under certain conditions, admittedly incalculable because of the inability to separate the product $k\epsilon$.

In addition to problems generated by insufficient data to fit the model properly, other factors have been postulated to account for observed or suspected anomalies in the reported values of k and ϵ . These include deviations from Beer's law,³ solvent interactions,^{4,5} the presence of complexes with other than 1:1 stoichiometry,⁵⁻⁷ and the use of more exact equations for numerical analysis.^{8,9} It is probable that many of these are apparent anomalies arising strictly from the incorrect use of the simple Benesi-Hildebrand¹⁰ or Scott¹¹ equations evaluating k and ϵ , especially when multiple equilibria might be a contributing factor. In view of the considerable diversity of interpretation of these anomalous results, it seems pertinent to reexamine some of these problems in greater detail. In the present paper, we will concern ourselves mainly with a simple model, reserving a discussion of a more complicated system for the following paper.

Basic Relationships. The formation constant of a weak molecular complex is commonly determined from measurements in which a dilute component P, maintained at a fixed total concentration $[P_0]$, is "titrated" by the addition of a second component X with total concentration $[X_0]$. When a 1:1 complex PX is formed to the exclusion of all other complexes, the formation constant is defined by the usual mass law expression

$$[PX] = k[P][X] = k([P_0] - [PX])([X_0] - [PX]) \quad (1)$$

and the entire experiment as $[PX]$ varies from zero to $[P_0]$ can be conveniently described in terms of the saturation fraction of the dilute component.

$$s \equiv [PX]/[P_0] = k[X]/(1 + k[X]) \quad 0 \leq s \leq 1 \quad (2)$$

The saturation fraction thus defined is identical, in the simple case, with the probability of binding,¹² or 1 minus the degree of dissociation of the complex.¹³

(2) Although not expressly stated, Person's argument applies as well to the separation of the product $k\delta_0$ where δ_0 is the chemical shift of the pure complex in nmr measurements, or to any other similar type of analysis.

(3) P. H. Emslie, R. Foster, C. A. Fyfe, and I. Horman, *Tetrahedron*, **21**, 2843 (1965).

(4) S. Carter, J. N. Murrell, and E. J. Rosch, *J. Chem. Soc.*, 2048 (1965).

(5) S. Carter, *ibid.*, **A**, 404 (1968).

(6) S. D. Ross and M. M. Labes, *J. Am. Chem. Soc.*, **79**, 76 (1957).

(7) G. D. Johnson and R. E. Bowen, *ibid.*, **87**, 1655 (1965).

(8) N. J. Rose and R. S. Drago, *ibid.*, **81**, 6138, 6141 (1959).

(9) G. Briegleb, "Elektronen-Donator-Acceptor-Komplexe," Springer-Verlag, Berlin, 1961.

(10) H. A. Benesi and J. H. Hildebrand, *J. Am. Chem. Soc.*, **71**, 2703 (1949).

(11) R. L. Scott, *Rec. Trav. Chim.*, **75**, 787 (1956).

(12) G. Weber in "Molecular Biophysics," B. Pullman and M. Weissbluth, Ed., Academic Press, New York, N. Y., 1965.

(13) F. J. C. Rossotti and H. Rossotti, "The Determination of Stability Constants," McGraw-Hill Book Co., Inc., New York, N. Y., 1961.

Equation 2 is the absorption isotherm, and the secondary concentration variable s is the essential parameter underlying Person's argument¹ concerning the minimum concentrations required to determine meaningful values of k . Person's limits for "region II" of $0.1/k$ and $9/k$ correspond to saturation fractions of 0.09 and 0.9, respectively, and the half-saturation point ($s = 0.5$) occurs at $[X] = 1/k$.

Two distinct experimental situations arise from the use of eq 2 as it stands, which have been discussed by Rose and Drago⁸ and by Briegleb.⁹ When the concentration of the free (unbound) excess component $[X]$ is of the same order of magnitude as the concentration of the complex, the substitution $[X] = [X_0] - [PX]$ is required,¹⁴ and eq 2 becomes quadratic in s (or $[PX]$). Conditions are easily arranged, however, such that the concentration of complex is negligible compared to the concentration of the excess component, in which case the substitution $[X] \simeq [X_0]$ can be made within a specified error. This error is

$$(k' - k)/k = -s[P_0]/[X_0] \quad (3)$$

where k' is the value calculated using the approximation $[X] \simeq [X_0]$, and k the value calculated from the exact solution $[X] = [X_0] - [PX]$. The error in k from this source can never exceed $[P_0]/[X_0]$ at any value of s , and is always negative; that is, k calculated from the approximation is always smaller than the true k . Because of the unidirectionality of this error, it is probably insufficient to neglect it when it merely falls within the experimental error of the measurements,⁹ except where the mean value is not appreciably changed. Careful distinction between this type of error and the statistical error arising from the measurements themselves should be maintained.

The saturation fraction is usually derived from measurements of some intensive parameter directly proportional to the concentration of the complex.¹³ An appropriate proportionality constant describes this correspondence; for optical absorption methods this is simply a statement of Beer's law for the complex PX. If A is the absorbance and ϵ the proportionality constant, $A = \epsilon[PX]$, and when all of the dilute component is complexed, the absorbance due to complex will have a maximum value $A_{\max} = \epsilon[P_0]$. The saturation fraction is then

$$A/A_{\max} = A/\epsilon[P_0] = [PX]/[P_0] \equiv s \quad (4)$$

Combination of eq 4 and eq 2 leads directly to the Benesi-Hildebrand¹⁰ or Scott¹¹ equations. The situation for nmr spectroscopy is similar, but in this case the observed chemical shift δ is proportional to the ratio of PX to P, and the proportionality constant is the chemical shift of pure complex δ_0 .

$$\delta = \delta_0[PX]/[P_0] \quad \delta/\delta_0 \equiv s \quad (5)$$

Combination of eq 5 with eq 2 yields the equation used by Hanna and Ashbaugh¹⁵ for nmr measurements of weak complex formation.

(14) This limitation does not apply to methods wherein a phase separation can be effected between $PX + P$ and X (equilibrium dialysis), or PX and $P + X$ (insoluble complex formation) since $[X]$ in this case can be obtained directly.

(15) M. W. Hanna and A. L. Ashbaugh, *J. Phys. Chem.*, **68**, 811 (1964).

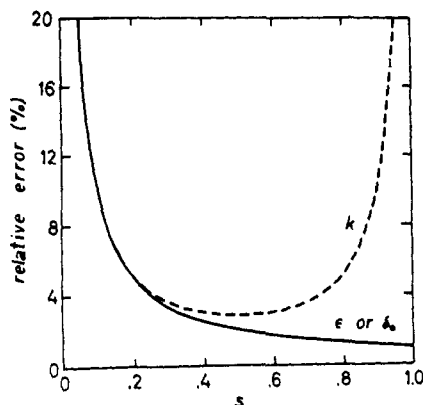


Figure 1. Per cent relative errors $100\Delta k/k$ and $100\Delta\epsilon/\epsilon = 100\Delta\delta_0/\delta_0$ as a function of the saturation fraction s . The curves are given for $\Delta s = 0.01$, which is equivalent to plotting the reduced relative errors $\Delta k/k\Delta s$ and $\Delta\epsilon/\epsilon\Delta s = \Delta\delta_0/\delta_0\Delta s$.

The basic problem rests with the simultaneous determination of two constants. When k is large and solubility permits, ϵ or δ_0 are easily determined independently under conditions where P is completely complexed, and the assumption of linearity of the intensive parameter with the concentration of complex can then be tested by dilution of the system under similar conditions. In practice, however, this is rarely possible, and both k , and ϵ or δ_0 must be determined from paired data or by curve fitting. The consequence of the latter limitation is that both the assumption of linearity of the intensive parameter with concentration and the assumption of 1:1 complex formation must be simultaneously verified. A failure of either assumption always has as its basis some sort of molecular aggregation phenomena (self-association, interaction with solvent molecules,^{4,5} or the formation of higher order complexes⁵⁻⁷) and phenomenological treatments based on deviations from either assumption separately³ are redundant since one implies the other. Where applicable, failure to control activity coefficients^{9,13} may also lead to apparent linearity or nonlinearity, as the case may be; this is merely another expression of the need to consider molecular aggregation phenomena.

Errors in k and ϵ or δ_0 . The argument put forward by Person¹ concerning the regions of the saturation curve in which k and ϵ (or δ_0) are not separable can be formalized in terms of the theoretical minimum errors in the various parameters. Applying the formula for the propagation of errors to the definition of k , Weber¹² has shown that if the errors in $[P_0]$ and $[X_0]$ are negligible compared to the error in s (or $[PX]$), the relative error in k is practically determined by the quantity

$$\frac{\Delta k}{k} \geq \Delta s \left[\frac{1}{s^2} + \frac{1}{(1-s)^2} \right]^{1/2} \quad (6)$$

Similarly, it can be shown that the errors in ϵ and δ_0 in terms of s are approximately given by

$$\Delta\epsilon/\epsilon = \delta_0/\delta_0 \geq \Delta s \sqrt{2}/s \quad (7)$$

again neglecting errors in $[X_0]$ and $[P_0]$, and using the approximation $\Delta s/s \simeq \Delta A/A$ or $\Delta\delta/\delta$. It should be stressed that the errors given by the equality in eq 6 and 7 are minimum values, and the obtention of these

minimum values in an actual experiment is difficult to achieve. The relative errors given by the equality in eq 6 and 7 are plotted as a function of s in Figure 1, from which the basis of Person's limit criteria are immediately obvious. The most accurate values of ϵ or δ_0 are obtained when the concentration of complex is maximal ($s = 1$), but reasonable accuracy can be expected above a saturation fraction of 0.2 or so. For k , the most accurate values will be obtained when the saturation fraction lies between 0.2 and 0.8; outside this range the determined values will become extremely uncertain.¹² In the type of complex formation measurements under consideration, k and ϵ are linked, and the "limits" for accurate simultaneous determination of both constants are approximately $0.2 \leq s \leq 0.8$.

Although Person's conclusions concerning the errors in k are in essential agreement with the error analysis presented by Weber,¹² the criterion that the highest concentration of the excess component be not less than approximately $0.1/k$ is insufficient. It is clear from Figure 1 that experimental points obtained as s decreases from $s = 0.2$ ($[X] = 0.25/k$) will have progressively less meaning in the calculation of k or ϵ because of their large errors. In an experiment performed between the limits $s = 0.1$ and $s = 0.01$, for example, the relative error in both k and ϵ will vary between ± 10 and $\pm 100\%$, respectively (with $\Delta s = 0.01$), and attempts to obtain accurate values of either parameter from the slope and intercept of a suitable plot can only end poorly. Thus the assignment of a limit criteria in which only the highest concentration of the excess component falls within the range of acceptable values is untenable.

Much more important, however, is the fact that the line obtained in the region $s < 0.1$ (or in any other limited region) may not be representative of the entire experiment. By choosing a relatively narrow range of values of s , nearly any function, no matter how wild, can be made to fit a "straight line," and the obtention of such a line over a limited portion of the saturation curves does not constitute a proof of the model. For a weak complex to fit the model "1:1 binding," for example, requires that the appropriate linear plot actually be linear over the entire range of values $0 \leq s \leq 1$. The point is that *measurements of a large portion of the saturation curve are required in order to assign experimental meaning to the slopes or intercepts of the fitted data*. The strict application of Person's limit criteria are to the line itself, not to the assignment of values for k and ϵ . In other words, a saturation fraction of 0.1 or so must be measured before it is certain that the fitted data have a finite slope or intercept within experimental error; this has no bearing on the assignment of k or ϵ unless the model is already precisely known for other considerations. In the case of multiple equilibria, for example,¹⁶ straight lines are invariably obtained at small saturation fractions, but the slopes and intercepts of such lines have widely differing experimental meaning.

A result often obtained from statistical analysis of experiments over a very small saturation fraction range ($s < 0.1$) is that the intercept (or slope, whichever corresponds to the traditionally assigned value " $k\epsilon$ ")

(16) See part II: D. A. Deranleau, *J. Am. Chem. Soc.*, **91**, 4050 (1969).

appears to have exceptional accuracy, while the slope (intercept) is largely in error. This is in reality no different from what we have been discussing: values obtained near or at the intercept will be accurately defined in terms of the line, although the slope has little or no meaning, and the assignment of given stoichiometry is impossible. When the entire saturation curve is measured, the errors in the slope or intercept have roughly the same statistical error. We note that the concept of sharpness of fit as used by Conrow, Johnson, and Bowen¹⁷ is hardly applicable to very limited regions of the saturation curve, especially below $s \sim 0.1$ or so, where with the exception of strong cooperative effects, virtually all binding data fit straight lines.¹⁶ In addition, correct application of normal linear regression analysis assumes *a priori* that the data are equally reliable and more or less evenly distributed over the entire range of values available to the chosen model. When these conditions are satisfied, the fitted data are a statement of the correspondence between the equation of the line and the equation of the model. The use of linear regression analysis to linearly fit part of the data similarly describes only the correspondence of that part of the equation of the model which happens to fit a straight line. In the case of measurements below $s = 0.1$, this is commonly interpreted as meaning that the other 90% of the data also fit the same line, a hazardous extrapolation at best.

The use of a second method to obtain estimates of k , such as nmr spectroscopy, is practical in those cases where limited solubility precludes measurements of the entire saturation curve. Under identical conditions, the optical and nmr methods should give similar results for the *apparent* k , provided that a valid slope or intercept can be measured. When different values are obtained, however, especially under different conditions, either (a) insufficient data are available to obtain a valid slope or intercept, and/or (b) the model chosen to represent the system (*e.g.*, "1:1 binding") is invalid. In any case, it is incorrect to assume³ that a better value for the extinction coefficient is obtained by division of the optical product $k\epsilon$ by the nmr k , since exactly the same considerations apply to the separation of the product $k\delta_0$ as to $k\epsilon$; that is, the product may well be accurately determined, but either value alone may be largely in error.

Criteria for the Proof of Fit of a Given Model. It is more or less obvious that measurements of a large portion of the saturation curve are required to show that the values k and ϵ or δ_0 calculated from a suitable linear plot have experimental significance in terms of a given model. The remaining question is *how much* of the theoretically obtainable data are necessary before the model, or range of models fitting the same phenomenological equation, can be stated with some confidence. Since this is fundamentally a problem of acquisition of information, it is logical to apply the principles of information theory to the binding process itself.¹²

The information obtained for several probabilities p_j , M in number, is¹⁵

$$I(p) = -K \sum_{j=1}^M p_j \ln p_j \quad \sum_{j=1}^M p_j = 1 \quad (8)$$

(17) K. Conrow, G. D. Johnson, and R. E. Bowen, *J. Am. Chem. Soc.*, **86**, 1025 (1964).

(18) L. Brillouin, "Science and Information Theory," Academic Press, New York, N. Y., 1962.

where $I(p) = i/N$ is the information per unit symbol for N symbols with average information i . When $K = (\ln 2)^{-1}$, $I(p)$ is given in "bits." For complex formation there are two probabilities p_j , binding and no binding;¹² hence, $M = 2$ in eq 8 and since we have previously defined the saturation fraction to be identical with the probability of binding (eq 2), $p_1 = s$, $p_2 = 1 - s$, and

$$I(s) = -K[s \ln s + (1 - s) \ln (1 - s)] \quad (9)$$

The saturation fraction scale can be divided into N identical segments of arbitrarily small length ds ; a particular value of the saturation fraction is then redefined as

$$s_i = n_i/N = n_i ds \quad (10)$$

and if $I(s_i)$ is the amount of information at a particular s_i , the maximum amount of information theoretically obtainable¹² is clearly the sum over the entire range of s .

$$I_{\max} = \sum_{i=1}^N I(s_i) = \int_0^1 I(s) ds \quad (11)$$

Similarly, the information at an intermediate value is

$$I = \sum_{i=1}^{n_i} I(s_i) = \int_0^{s_i} I(s) ds \quad (12)$$

and the fraction of the total theoretically available information accumulated as s varies from zero to a particular value s_i is, from eq 9, 11, and 12

$$I/I_{\max} = s - s^2 \ln s + (1 - s)^2 \ln (1 - s) \quad (13)$$

The rate of accumulation of information can be found by normalizing eq 13 with respect to s , obtaining

$$I/sI_{\max} = 1 + s \ln s + s^{-1}(1 - s)^2 \ln (1 - s) \quad (14)$$

which has a maximum corresponding to 83% of the total possible information at $s = 0.76$. This point represents the optimum in the accumulation of information concerning the fundamental binding processes, and we would like to suggest that *measurements should be made over roughly 75% of the saturation curve before the data can be considered as truly representative of the entire experiment.* The functions represented by eq 9, 13, and 14 are shown in Figure 2. It should be noted in particular that the fractional information is less than the corresponding saturation fraction below $s = 0.5$; for an experiment in which data are obtained over the region $0 \leq s \leq 0.1$ (Person's region I¹), the amount of information accumulated is only 4% of the total information obtainable, even though 10% of the saturation curve has been measured. We also note that increasing the number of measurements made in the range $0 \leq s \leq 0.1$ does not change this situation, as we are dealing with the fraction of theoretically accumulated to potentially available information. The ratio I/I_{\max} is independent of the number of measurements, which we have purposely not specified in order to perform the integrations in eq 11 and 12. However, if the number of points actually measured is less than the number theoretically possible (limited by the condition¹² $ds \leq 2\Delta s$), the ratio of *actually* accumulated to theoretically possible information can only decrease relative to I/I_{\max} as defined above. Equation 13 as derived therefore represents the upper

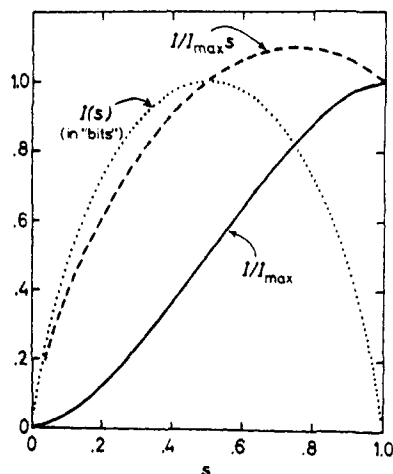


Figure 2. Information $I(s)$ as a function of s (dotted curve); accumulated information relative to total possible information (I/I_{\max}) as s varies from $0 \rightarrow 1$ (solid curve); and rate of accumulation of information ($I/I_{\max}s$) as s varies from $0 \rightarrow 1$ (dashed curve); see text for details.

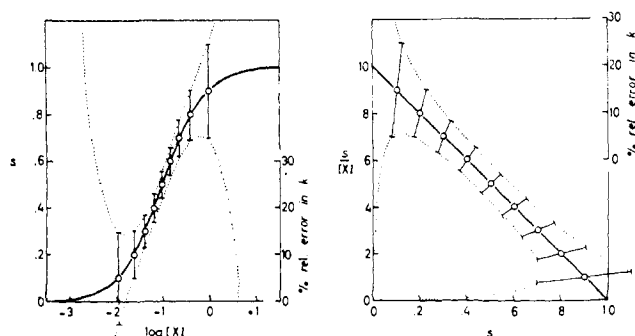


Figure 3. Bjerrum's formation function (left) and the Scatchard plot (right), in s units. To obtain the curves in the same units as given in the text (eq 15 and 16), substitute $A/[P_0]$ or δ for s , and multiply the units of the axes containing s by ϵ . The limits, intercepts, and slope are then given directly by eq 15 or 16. The curves were calculated for $k = 10$, and points are shown for equal increments in s . The error was calculated according to eq 6 with $\Delta s = 0.01$. Note the error is symmetrical with respect to the origin on the Scatchard plot (C. de Haën, personal communication).

limit of possible accumulation of information, and the actual experimental situation is usually much worse.

Choice of Plot and Errors Arising from the Plot. Since it is not, in general, possible to calculate the saturation fraction for weak complexes without prior knowledge of k , the choice of a suitable means of graphical presentation is of importance; in fact it is *only* by such means that the range of saturation fraction being studied can be easily determined. Of the several types of plots resulting from rearrangement of eq 2 in one of its substituted forms, only three are suitable means of data presentation.^{12, 19, 20} These are: (a) direct plot of $A/[P_0]$ (or δ) vs. $[X]$ (absorption isotherm, Figure 1 of Person¹) according to the observation equations

$$\frac{A}{[P_0]} = \frac{k\epsilon[X]}{1 + k[X]} \text{ or } \delta = \frac{k\delta_0[X]}{1 + k[X]} \quad (15)$$

(19) B. H. J. Hofstee, *Science*, **116**, 329 (1952).

(20) J. T. Edsall and J. Wyman, "Biophysical Chemistry," Vol. I, Academic Press, New York, N. Y., 1958.

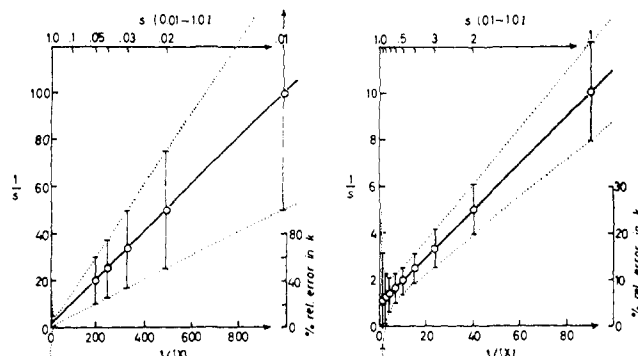


Figure 4. The Benesi-Hildebrand, or reciprocal plot (in s units) showing two overlapping regions of the saturation fraction. To obtain the usual form of the plot for charge-transfer calculations, substitute $A/[P_0]$ or δ for s and multiply the units of the $1/s$ axis by $1/\epsilon$. Points and error bars are as in Figure 3, and $k = 10$. Note the difference in per cent relative error scales.

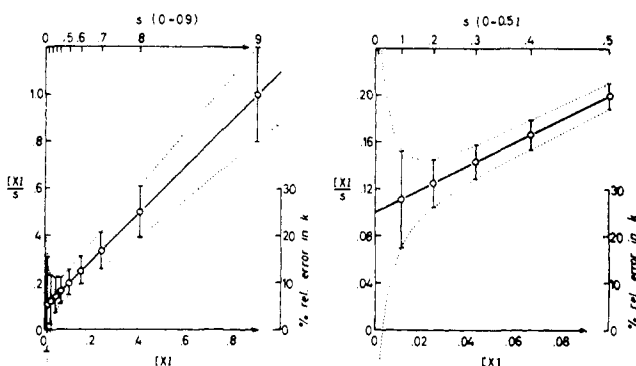


Figure 5. The Scott, or half-reciprocal plot (in s units), showing two overlapping regions of the saturation curve. To obtain the usual form of the plot for charge-transfer calculations, substitute $A/[P_0]$ or δ for s and multiply the units of the $[X]/s$ axis by $1/\epsilon$. Points and error bars are as in Figure 3, and $k = 10$.

(b) plot of $A/[P_0]$ (or δ) vs. $\log [X]$ (Bjerrum's formation function²¹ or the "titration" curve, Figure 3); (c) plot of $A/[P_0][X]$ vs. $A/[P_0]$, or $\delta/[X]$ vs. δ (Scatchard plot,²² Figure 3) according to the equations

$$\frac{A}{[P_0][X]} = k\left(\epsilon - \frac{A}{[P_0]}\right) \text{ or } \frac{\delta}{[X]} = k(\delta_0 - \delta) \quad (16)$$

A satisfactory plot is one on which all of the theoretically obtainable data can be plotted, whether experimentally obtainable or not, and this criterion is satisfied by either the formation function curve or the Scatchard plot (Figure 3), the former because of the logarithmic scale and the latter because the abscissa varies linearly with the saturation fraction. The direct absorption isotherm (rectangular hyperbola), the Benesi-Hildebrand or reciprocal plot (Figure 4), and the Scott or half-reciprocal plot (Figure 5) all have open upper limits on the abscissa, and the latter two are particularly objectionable because they are *not* curved. This means that by either of these two plots, any range of values of the saturation fraction can be represented as a complete plot. With the formation function curve or the Scatchard plot, however, the experimenter is not free to choose the scale on which the data are

(21) J. Bjerrum, "Metal Ammine Formation in Aqueous Solution," P. Hase and Son, Copenhagen, 1941.

(22) G. Scatchard, *Ann. N. Y. Acad. Sci.*, **51**, 660 (1949).

plotted, as the proper scale is that which contains all of the saturation curve and is visually apparent on both these plots. Failure to recognize these facts has led to a large amount of confusion in the case of the Benesi-Hildebrand and Scott plots. As already discussed, the smaller the range of saturation fraction being studied, the better will be the possibility that a straight line is obtained from either of these plots even if the entire experiment is represented by a highly curved line, which we note specifically is a general characteristic of equilibria involving other than 1:1 stoichiometry.¹⁶ In addition, the Benesi-Hildebrand treatment, and to a lesser degree, the Scott treatment, unequally weigh points evenly spaced on the saturation fraction scale. For proper evaluation of k by linear regression techniques (least squares), the experimental points should be obtained for equal increments in s , and weighing factors should be applied to the linear regression equations for all data obtained otherwise.²³ In the case of the Benesi-Hildebrand treatment, this is particularly disastrous, as each addition of an equal increment in s results in an approximate doubling of the statistical weight factor. For six successive additions, for example, the most dilute (and least accurate) value will have a statistical weight factor nearly equal to that of the four most concentrated values. Linear regression analysis of this case without the introduction of weighing factors would essentially draw the line through only three points, including the two most dilute and least accurate values. We note that this "reciprocal doubling factor" also drastically reduces the chance of seeing curvature in the plot if it is actually present. The Scatchard plot has been used previously for charge-transfer calculations by Foster and Fyfe,²⁴ and if a linear plot is desired to fit the data, this plot is to be preferred. It should be pointed out, however, that an important advantage of this plot is that the quality of the fit is immediately apparent provided the point 0,0 (origin) is included in the graph, which Foster and Fyfe failed to do. The best method of obtaining the constants k and ϵ is to fit the observation equation (eq 15) directly, as has been clearly demonstrated in the important article by Birge.²³ This method gives

(23) R. T. Birge, *Rev. Mod. Phys.*, **19**, 298 (1947).

(24) R. Foster and C. A. Fyfe, *Trans. Faraday Soc.*, **61**, 1626 (1965).

comparable results for k and ϵ with no ambiguity in the computed errors, and although slightly more complicated than the straight-line fitting procedure, it has been successfully used previously for the computation of binding constants.²⁵

Conclusions

In summary, we have attempted to pinpoint various sources of error arising from the usual methods of treatment of weak molecular complexes, and to suggest further criteria for the reliability of formation constants and extinction coefficients so obtained. In particular, the obtention of roughly 75% of the data comprising the complete saturation curve seems necessary before the model can be considered proven by any single equilibrium technique. This conclusion is not restricted to weak molecular complexes, but is generally applicable to *all* cases of complex formation. Similar conclusions as those given by the application of information theory can be obtained in principle by considering the accuracy of the values of k and/or ϵ , since the accuracy α is related to the information through¹⁸

$$\Delta I = -K \ln \epsilon_c = K \ln \alpha$$

where ΔI is the amount of information obtained and ϵ_c is the comparative error.

Of secondary importance are the conclusions concerning the proper choice of plot, the error arising from the approximation $[X] = [X_0]$, and the comments on linear regression analysis. It should be reemphasized, however, that the limit criteria given by Person¹ apply not to the values " k ," " ϵ ," or " $k\epsilon$," but to the limiting slope and intercept of the line describing the smallest possible measurable range of saturation fractions for which finite, meaningful values can be obtained, and no interpretation in terms of a model can be made from such values alone.

Acknowledgment. This work was supported by the National Institutes of Health, and the Swiss National Foundation for Scientific Research, Grants No. 3426 and 4883. The author wishes to thank Ch. de Haën for many stimulating and interesting discussions concerning the content of this paper.

(25) D. Deranleau and H. Neurath, *Biochemistry*, **5**, 1413 (1966).