Equilibria in Systems with Polynuclear Complex Formation

III. Derivation of a Generalized Bodländer Formula and a General Method of Calculating Equilibrium Constants

BENGT HEDSTRÖM

Department of Inorganic Chemistry, Royal Institute of Technology, Stockholm 70, Sweden *

A number of relationships have been derived which apply to equilibrium mixtures of two reactants, A and B, and their condensation products, of the general type A_pB_q . Denoting total, stoichiometric concentrations with capital letters, and equilibrium concentrations with small letters, the following relationship, concisely stated in Jacobian notation, is found to be universally valid:

$$J(B, \ln b / \ln a, A) \equiv \frac{\partial(B, \ln b)}{\partial(\ln a, A)} = 1$$

Solving for $a = a/A$, the following equation results

$$\log a = \left[\int_0^B (\partial \log \beta / \partial A)_B dB \right]_A$$

From this equation, a and hence a , can be calculated graphically from a set of known values of A , B and $\beta = b/B$. The equations presuppose constant activity factors, an assumption which is approximately correct for solutions of high, and constant ionic strength.

In the latter part of the paper a straightforward method is presented which can be used to find the composition of poly- and mononuclear complexes and the corresponding equilibrium constants from experimental data on any set of equilibria between two reactants.

The methods presented in this paper have been successfully applied by the author in a study on the hydrolysis of the iron(III) ion, Fe^{3+} .

This paper is number three in a series dealing with the problem of interpreting experimental data on polynuclear complex formation. In the two previous papers (AB I¹, AB II²) Sillén has worked out in detail the mathematics of so-called "core and links" mechanisms. He derived very useful and simple criteria by means of which data on polynuclear complex formation can be tested. Although the methods developed in AB I and AB II seem to be of far-

* Present address: Mo och Domsjö AB, Husums Sulfatfabrik, Husum, Sweden.

reaching applicability, they do not cover all possible types of complex formation. Sillén purposely restricted his investigation to the "core and links" type of complex, which he has found to be rather common.

The field of polynuclear complex chemistry is in its infancy and it is, therefore, difficult to know which of several possible methods of interpreting experimental data will be the most useful in the majority of cases met with in practice. The accuracy of experimental data is, indeed, limited. One may therefore question if it is at all possible, at the present time, to prove the existence of more than a very limited number of species in a complex forming system, or, to put it another way, one may ask if experimental data, even in a favorable case, can supply more information than that represented by numerical values of, say three or four equilibrium constants. If this is so, it might, in an actual case, be very difficult to know exactly what is proof of the existence of the different possible complexes and what is merely a convenient method of correlating the experimental data in terms of the law of mass action. The methods in AB I and AB II should in many cases be useful for rejecting faulty hypotheses and also in choosing the one which is likely to be most consistent with the measurements.

In this paper a somewhat different method of attack is used in an effort to develop, in principle, a generally valid method of interpreting data on the formation of polynuclear complexes of the general type A_pB_q . The leading principle in the present work can be stated in the following way: No *a priori* assumption, whatsoever, is made as to the compositions or abundance of the complexes formed. This implies that the compositions of the complexes and the corresponding equilibrium constants will be obtained directly from a straight-forward method of calculation. In the author's opinion, the methods here presented should be of value as a supplement or an alternative to the methods in AB I and AB II.

The first part of the paper is devoted to the problem of calculating the equilibrium concentration of one of the reactants from a series of known values of the equilibrium concentration of the other and the total concentrations of the two reactants. Very often it is not possible to measure the equilibrium concentrations of both reactants directly. It is obvious that in such a case the interpretation of the data is made simpler if the unknown concentration can first be calculated, without presupposing any particular reaction mechanism.

In the second part of the paper, a method of interpretation of equilibrium data is presented which should be applicable to any type of mono- or polynuclear complex system, *provided data of sufficient accuracy are available*.

This paper is mathematical in character. The author has purposely omitted detailed discussions or examples of applications in order that the style should be as concise and as simple as possible. The methods presented here have been successfully applied in the author's investigation on the hydrolysis of the iron(III) ion, Fe^{3+} . The reader is referred to this work³ for a detailed example of the application of the methods given in this paper.

List of symbols

- A = total, stoichiometric concentration of the reactant A
 a = equilibrium concentration of A
 B = total, stoichiometric concentration of the reactant B
 b = equilibrium concentration of B
 $F_{q+1} = (F_q - I_q)/ab$
 $I_q = q \sum_p \kappa_{pq} a^{p-1}$
 p = number of A atoms in the complex $A_p B_q$
 q = number of B atoms in the complex $A_p B_q$
 S = complexity sum (4)
 x, y = arbitrary, independent variables, used in equations 10 and 10 a (see text)
 $\alpha = a/A$
 $\beta = b/B$
 κ_{pq} = equilibrium constant for the reaction $pA + qB = A_p B_q$

DERIVATION OF A GENERALIZED BODLÄNDER'S FORMULA

If possible, emf-methods are generally preferred in the study of complex formation. The reasons are that they are accurate and relatively simple from the experimental point of view. A few recent modifications of the emf-method for studying complex equilibria have been of great importance in simplifying both the experimental technique and the interpretation of the data obtained.

These modifications are: Activity factors are kept constant by working with solutions of high, constant ionic strength. The equilibrium solutions under study are prepared by a titration technique⁴. In this way a large amount of experimental data can be obtained in a relatively short time and with a minimum of manipulation. By using two (in some cases three) burets during an emf-titration, the total, stoichiometric concentration of one of the reactants can be kept constant, while the other is varied. A series of such titrations at different total, stoichiometric concentrations of one of the reactants often covers the whole range of complexes of the two reactants, and is very convenient for the mathematical interpretation of the data.

The aim of the following reasoning and deductions is to show how the equilibrium concentration of one of the reactants is fixed by the other in all circumstances, and how it can be calculated from data of the kind mentioned above, without presupposing any particular reaction mechanism.

In the case of a *mononuclear* process the solution of this problem is the well-known Bodländer formula⁴ (see eq. 14 below).

The formulas derived below may be considered as extensions of this formula. In contrast to Bodländer's formula they are, however, applicable to any set of mono- and *polynuclear* complexes.

Consider an equilibrium involving two reactants A and B, forming a condensation product $A_p B_q$:

In the special case of an equilibrium between a metal ion and a complex forming substance, B may symbolize the metal ion and A the ligand. A set

of equilibria of this kind may exist simultaneously, since p and q may, in general, have any positive, integral values.

On applying the law of mass action to these equilibria, equations of the following general type are obtained:

$$\frac{[A_p B_q]}{a^p b^q} = \kappa_{pq} \quad (1)$$

where κ_{pq} denotes the equilibrium constant, and a and b the equilibrium concentrations of A and B, respectively. Activity coefficients are assumed to be constant, as a consequence of working in a medium of high, and constant ionic strength; the law of mass action will therefore apply to concentrations as well as to activities. From stoichiometric considerations

$$A = a + \sum p \kappa_{pq} a^p b^q \quad (2)$$

$$B = b + \sum q \kappa_{pq} a^p b^q \quad (3)$$

where A and B denote total, stoichiometric concentrations. The summations have been carried out over all pairs of values of p and q except (1;0) and (0;1). The absence of competitive complex-formation with any third reagent C is assumed.

From now on the problem is a purely mathematical one. The method adopted by the author in order to find the universal relationship between a and b , consists of an elimination of the unknowns p , q , and κ_{pq} by successive, partial differentiation.

Let us first define the function S

$$S = \sum \kappa_{pq} a^p b^q \quad (4)$$

Sillén¹ introduced the term "complexity sum" for this expression. Partial differentiation of (4) with respect to $\ln a$ and $\ln b$ gives

$$\left(\frac{\partial S}{\partial \ln a} \right)_b = a \left(\frac{\partial S}{\partial a} \right)_b = \sum p \kappa_{pq} a^p b^q \quad (5)$$

$$\left(\frac{\partial S}{\partial \ln b} \right)_a = b \left(\frac{\partial S}{\partial b} \right)_a = \sum q \kappa_{pq} a^p b^q \quad (6)$$

Substituting into equations (2) and (3)

$$\left(\frac{\partial S}{\partial \ln a} \right)_b = A - a \quad (7)$$

$$\left(\frac{\partial S}{\partial \ln b} \right)_a = B - b \quad (8)$$

A second differentiation of (7) with respect to $\ln b$ and of (8) with respect to $\ln a$, and application of the identity

$$\left(\frac{\partial^2 S}{\partial \ln a \partial \ln b} \right) = \left(\frac{\partial^2 S}{\partial \ln b \partial \ln a} \right)$$

finally gives

$$\left(\frac{\partial B}{\partial \ln a} \right)_b = \left(\frac{\partial A}{\partial \ln b} \right)_a \quad (9)$$

At this point the unknowns p , q , and κ_{pq} have been eliminated.

Equation (9) is not yet very useful in the calculation of a from b , or *vice versa*, since a and b cannot be expressed explicitly. It may, however, be transformed to useful relationships by a transformation of variables. As in the case of the usual thermodynamic functions, the most convenient way of performing this transformation is by means of the Jacobian (functional) determinant^{5,6}. In terms of Jacobians, equation (9) can be written

$$\frac{\partial(B, \ln b)}{\partial(\ln a, \ln b)} = \frac{\partial(A, \ln a)}{\partial(\ln b, \ln a)} \quad (9a)$$

from which we get the most concise, basic form of the universal relationship between the variables involved

$$J(B, \ln b / \ln a, A) \equiv \frac{\partial(B, \ln b)}{\partial(\ln a, A)} = 1 \quad (9b)$$

If two arbitrary quantities x and y are chosen as independent variables, it follows from (9b)

$$\frac{\partial(B, \ln b)}{\partial(x, y)} = \frac{\partial(\ln a, A)}{\partial(x, y)} \quad (10)$$

or in determinant notation

$$\begin{vmatrix} \left(\frac{\partial B}{\partial x}\right)_y & \left(\frac{\partial B}{\partial y}\right)_x \\ \left(\frac{\partial \ln b}{\partial x}\right)_y & \left(\frac{\partial \ln b}{\partial y}\right)_x \end{vmatrix} = \begin{vmatrix} \left(\frac{\partial \ln a}{\partial x}\right)_y & \left(\frac{\partial \ln a}{\partial y}\right)_x \\ \left(\frac{\partial A}{\partial x}\right)_y & \left(\frac{\partial A}{\partial y}\right)_x \end{vmatrix} \quad (10a)$$

As an illustration, let $x = \ln a$ and $y = \ln b$. We then get from equation (10)

$$\left(\frac{\partial B}{\partial \ln a}\right)_b = \left(\frac{\partial A}{\partial \ln b}\right)_a$$

which is identical with equation (9). If we focus attention on the five quantities A , B , $\ln a$, $\ln b$, and $\beta = b/B$, ten pairs of independent variables are possible. For illustrative purposes a few of these cases have been worked out below. It is a simple matter to derive formulas based upon other sets of independent variables.

Special choices of independent variables

With A and B as independent variables, (10) takes the form

$$\left(\frac{\partial \ln b}{\partial A}\right)_B = \left(\frac{\partial \ln a}{\partial B}\right)_A \quad (11)$$

On introducing $\alpha = a/A$, and $\beta = b/B$, instead of a and b , we get from (11)

$$\left(\frac{\partial \ln \beta}{\partial A}\right)_B = \left(\frac{\partial \ln \alpha}{\partial B}\right)_A \quad (11a)$$

With B and $\ln a$ as independent variables (10) gives

$$\left(\frac{\partial \ln b}{\partial \ln a}\right)_B = -\left(\frac{\partial A}{\partial B}\right)_a \quad (12)$$

If we assume that all complexes are *mononuclear in B*, thus can be written as BA_n , then

$$S = b \sum n_n a^n = b \cdot f(a) \quad (13)$$

and from (7) and (8)

$$\begin{aligned} A &= a + baf'(a); & B &= b + bf(a) \\ \left(\frac{\partial A}{\partial B}\right)_a &= \frac{af'(a)}{1+f(a)} = \frac{A-a}{B} \end{aligned} \quad (13a)$$

For complexes mononuclear in B , (12) thus takes the form

$$\left(\frac{\partial \ln b}{\partial \ln a}\right)_B = -\frac{A-a}{B} \quad (14)$$

which is identical with Bodländer's well-known formula; $\frac{A-a}{B}$ is the average ligand number, which is often denoted by $\bar{n}$; in AB I and AB II it was denoted Z .

Choosing A and $\ln b$ as independent variables (10) gives an equation of the same form as (12)

$$\left(\frac{\partial \ln a}{\partial \ln b}\right)_A = -\left(\frac{\partial B}{\partial A}\right)_b \quad (15)$$

With B and $\ln b$, we find

$$\left(\frac{\partial \ln a}{\partial B}\right)_b \left(\frac{\partial A}{\partial \ln b}\right)_B - \left(\frac{\partial A}{\partial B}\right)_b \left(\frac{\partial \ln a}{\partial \ln b}\right)_B = 1 \quad (16)$$

With B and β as independent variables, (10a) yields

$$\left(\frac{\partial A}{\partial \ln \beta}\right)_B \left(\frac{\partial \ln a}{\partial B}\right)_\beta - \left(\frac{\partial \ln a}{\partial \ln \beta}\right)_B \left(\frac{\partial A}{\partial B}\right)_\beta = 1 \quad (17)$$

Choosing $\ln a$ and β as independent variables, one gets

$$\left(\frac{\partial A}{\partial \beta}\right)_a + \left(\frac{\partial \ln b}{\partial \ln a}\right)_\beta \left(\frac{\partial B}{\partial \beta}\right)_a - \left(\frac{\partial B}{\partial \ln a}\right)_\beta \left(\frac{\partial \ln b}{\partial \beta}\right)_a = 0 \quad (18)$$

Finally with $\ln b$ and β as independent variables, one gets from (10a)

$$\left(\frac{\partial B}{\partial \beta}\right)_b - \left(\frac{\partial \ln a}{\partial \beta}\right)_b \left(\frac{\partial A}{\partial \ln b}\right)_\beta + \left(\frac{\partial A}{\partial B}\right)_b \left(\frac{\partial \ln a}{\partial \ln b}\right)_\beta = 0 \quad (19)$$

DISCUSSION

It should be pointed out that the equations derived in the preceeding paragraph apply to any equilibrium mixture of the two reactants A and B, forming one or more condensation products of the general type A_pB_q . A and B may, for instance, be two organic compounds, or hydroxyl ion and a metal ion, forming a series of mono- and polynuclear complexes.

Usually one knows from experiments at least three of the four variables involved, for instance A, B, and b. One may then wish to find a also, before proceeding to interpret the data in terms of equilibrium constants for the a priori unknown equilibria.

Equation (11), or its equivalent (11a), seems to be most suitable to the calculation of a, since it can be expressed explicitly in terms of the other variables by this equation.

As equation (11) is a partial differential equation, it is satisfied by an infinite number of functions b (A, B, a) or a (A, B, b). The function a sought is the one which satisfies both (11) and the boundary conditions, as given by the experimental data. There are at least two generally valid boundary conditions

$$\ln a = 0, \text{ for } B = 0 \quad (20a)$$

and

$$\ln \beta = 0, \text{ for } A = 0 \quad (20b)$$

which are independent of the actual reaction mechanism. On account of the symmetrical nature of (11), we shall consider only the first one. Solving equation (11a) for log a gives

$$\log a = \left[\int_0^B \left(\frac{\partial \log \beta}{\partial A} \right)_B dB \right]_A \quad \text{or} \quad \log \frac{a}{A} = \left[\int_0^B \left(\frac{\partial \log \frac{b}{B}}{\partial A} \right)_B dB \right]_A \quad (21a)$$

This relation suggests a relatively simple and straightforward method of calculating log a. For each series of experiments, at constant B, one plots log β against A, determines the derivatives of these curves at a certain A value, and finally the corresponding value of log a from the area under $(\partial \log \beta / \partial A)_B$, when plotted against B.

Conversely, one may also calculate b from data of the form (A, B, a), applying the equation

$$\log \beta = \left[\int_0^A \left(\frac{\partial \log a}{\partial B} \right)_A dA \right]_B \quad (21b)$$

The graphical method discussed here suggests that measurements at a relatively large number of different values of B (or of A, respectively), including low values, are desirable in the calculation of a from b (or *vice versa*). This is one reason for trying to extend the range of concentrations as much as possible in investigating polynuclear equilibria.

For further details on the use of (21), the reader is referred to the author's paper on the hydrolysis of the iron(III) ion³.

EVALUATION OF EQUILIBRIUM CONSTANTS OF A POLYNUCLEAR COMPLEX FORMATION PROCESS

In the evaluation of the equilibrium constants of a polynuclear process a knowledge of both a and b is highly desirable. If, therefore, only one of the two equilibrium concentrations can be measured directly, the other may be calculated by the methods developed in the preceding section, which presupposes no particular reaction mechanism. When both a and b are known, the equilibrium constants may be calculated in a rather straightforward manner, as will now be shown.

Equation (3) may be written in the following way

$$F_1 = \frac{(B - b)}{ab} = \sum_{p,q} q \kappa_{pq} a^{p-1} b^{q-1}$$

When $q = 1$, then F_1 will be a function of a only, independent of b . Bjerrum⁷, Leden⁴ and others have already given detailed accounts of the problem of dealing with mononuclear processes.

In order to evaluate the "degree" ($q_{\max}$) of the polynuclear process, we may plot F_1 against b at a certain constant value of a . If a straight line with an intercept, I_1 , and a finite slope is obtained, then only mono- and dinuclear complexes are present; if a parabola, trinuclear complexes are also present, and so on. It may be difficult to decide between curves of higher degrees than the first, but this problem can be solved in the following way.

From the plot of F_1 against b the intercept

$$I_1 = \sum_p \kappa_{p1} a^{p-1}$$

is obtained. If then

$$F_2 = \frac{(F_1 - I_1)}{ab}$$

is plotted against b , the new intercept is given by

$$I_2 = 2 \sum_p \kappa_{p2} a^{p-2}$$

This same procedure is repeated until a straight line is obtained. The "steps" ($q = 1, 2, 3, \dots$) of the process will be successively separated as intercepts:

$$I_q = q \sum_p \kappa_{pq} a^{p-q}$$

$$F_{q+1} = \frac{F_q - I_q}{ab}$$

For each value of q , p can, in the general case, assume more than one value. In order to find all equilibrium constants, one must compare the F -curves obtained at different, constant values of a . From the resulting system of equations all κ_{pq} 's may be calculated.

This method is laborious, but it appears to be the best method available, that does not involve any presuppositions as to the actual mechanism. In practice, the accuracy of the measurements is limited, and the data will probably be exhausted after the introduction of but a few constants. The calculations may therefore in reality become relatively simple. As an example, it may be mentioned that in the application of this method of calculation to the hydrolysis of Fe^{3+} , the data could be explained, within the limits of experimental error, by introducing only three constants, κ_{11} , κ_{12} and κ_{22} , corresponding to the formation of FeOH^{2+} , $\text{Fe}(\text{OH})_2^+$ and $\text{Fe}_2(\text{OH})_2^{4+}$, respectively.

On the other hand, it might happen that the actual reaction mechanism consists of a very large or even infinite number of steps ($q = 1 \dots \infty$), as has been suggested⁸ for the hydrolysis of Bi^{3+} and, more recently^{9, 10}, for Th^{4+} and UO_2^{2+} . This should be indicated by the fact that in such a case the data will not yield a straight line even for high orders (high q -values) of F_q , when plotted against a .

Acknowledgement. The author is much indebted to Professor L. G. Sillén for his advice and inspiring interest in the work.

Thanks are also due to Dr F. Rossotti, Oxford, for his valuable criticism of the paper.

This work is part of a program supported by *Atomkommittén* (Swedish Atomic Committee) and *Statens Naturvetenskapliga Forskningsråd* (Swedish Natural Science Research Council). The main part of it was performed while the author had a research fellowship from *Statens Tekniska Forskningsråd* (Swedish Technical Research Council).

REFERENCES

1. Sillén, L. G. *Acta Chem. Scand.* **8** (1954) 299 (AB I).
2. Sillén, L. G. *Acta Chem. Scand.* **8** (1954) 318 (AB II).
3. Hedström, B. *Arkiv Kemi* **6** (1953) 1.
4. Leden, I. *Dissertation*, Lund 1943.
5. Margenau, H. and Murphy, G. M. *The mathematics of physics and chemistry*, van Nostrand, New York 1943.
6. Sherwood, T. K. and Reed, C. E. *Applied mathematics in chemical engineering*, McGraw Hill, New York 1939.
7. Bjerrum, J. *Dissertation*, Copenhagen 1941.
8. Granér, F. and Sillén, L. G. *Acta Chem. Scand.* **1** (1947) 631.
9. Hietanen, S. *Acta Chem. Scand.* **8** (1954) 1607.
10. Åhrland, S., Hietanen, S. and Sillén, L. G. *Acta Chem. Scand.* **8** (1954) 1907.

Received January 25, 1955.

Note added in proof: Equation (21a), which was applied in 1952 to the iron(III) hydrolysis³ is related to eq. (24) in a paper of McKay (*Trans. Faraday Soc.* **49** (1953) 237). McKay's equation may be obtained from (12) whereas (21a) is derived from (11).