and

$$\frac{1}{c_2} = \frac{1}{m_2 d} + \frac{M_2}{1000d} + \frac{n_3 M_3}{1000 n_2 d}$$

Therefore

$$\frac{1}{m_2} = \frac{d}{c_2} - \frac{M_2}{1000} - \frac{n_3 M_3}{1000 n_2}$$

and

$$\frac{n_1 \bar{v}_1^\circ}{n_2} = \frac{1000}{d_0}\left[\frac{d}{c_2} - \frac{M_2}{1000} - \frac{n_3 M_3}{1000 n_2}\right]$$

Thus

$$\phi(V_2) = \frac{1000}{c_2} - \frac{1000d}{c_2 d_0} + \frac{M_2}{d_0} + \frac{n_3 M_3}{n_2 d_0} - \frac{n_3}{n_2}\phi(V_3)$$

$$= \frac{1000}{c_2}\left(\frac{d_0 - d}{d_0}\right) + \frac{M_2}{d_0} + \frac{n_3}{n_2}\left[\frac{M_3}{d_0} - \phi(V_3)\right]$$

EVANSTON, ILLINOIS

[CONTRIBUTION FROM OAK RIDGE NATIONAL LABORATORY AND THE LABORATORY OF BIOCHEMISTRY, NATIONAL CANCER INSTITUTE, NATIONAL INSTITUTES OF HEALTH, PUBLIC HEALTH SERVICE]

New Methods for the Calculation of Association Constants of Complex Ion Systems[1]

BY JOHN Z. HEARON AND JAMES B. GILBERT

RECEIVED JULY 30, 1954

Extensions of and additions to existing methods of computing association constants are presented. New methods, with special reference to data from potentiometric titration, are described in detail for the cases of one, two or three complexes. These methods give a calculated value for some, or all, of the constants for each point on the titration curve. A general equation for the concentration of free ligand is derived and the limits of applicability discussed. A new method for computing the concentration of free ligand and $\bar{n}$ from the known constants is discussed.

I. Introduction

In practice it is usually possible to determine directly, or to compute from observed quantities, the concentrations of only a few of the relevant chemical species in a complex ion system. Previous investigations[2-5] of the computation of association constants for such systems have produced no clear discussion of the number of known relations in such a system and no attempt to exploit the mathematical properties of these relations for the purpose of computing association constants in a more rigorous manner. The present paper is presented with that aim. In particular, there is considered the determination of association constants for an ampholyte–metal ion system on the basis of data from potentiometric titration. While we have in mind specifically the association of peptides and amino acids with metal ions[6] the formulation is fairly general. This work had been completed[1] when several relevant reports appeared (especially ref. 4, 5) and the relation of this investigation to certain others will be discussed.

II. Some General Considerations

Consider the equilibria

$$\varphi + M\varphi_{k-1} \rightleftarrows M\varphi_k,\ k = 1, 2, \ldots, \alpha \qquad (1)$$

where M, the central ion or molecule, is of valence $+p$, the ligand φ of valence $-m$, and the valence of the kth complex, $M\varphi_k$, is $p - km$. It is assumed that the ligand φ is the completely ionized form of φH_n and is formed in the nth of the equilibria

$$\varphi H_{n-k+1} \rightleftarrows \varphi H_{n-k} + H^+,\ k = 1, 2, \ldots, n \qquad (2)$$

Proton-bearing complexes (*e.g.*, $(\varphi H_j)M$, $j > 0$) and polynuclear complexes (*e.g.*, $M_j\varphi_k$, $j > 1$) are excluded from consideration. This limitation of the analysis to follow is discussed later. With this understanding, then, the "totality conditions" or material balances are

$$c = \sum_{k=0}^{n} [\varphi H_k] + \sum_{k=0}^{\alpha} k[M\varphi_k] \qquad (3)$$

$$b = \sum_{k=0}^{\alpha} [M\varphi_k] \qquad (4)$$

where c and b are the total concentration of ligand and metal. The condition of electroneutrality is

$$S + \sum_{k=0}^{n} (k - m)[\varphi H_k] + \sum_{k=0}^{\alpha} (p - mk)[M\varphi_k] = 0 \qquad (5)$$

where $S = \Sigma \nu_i [I_i]$, and ν_i and $[I_i]$ are the valence and concentration of the ith ion which contains neither M nor φ. We assume S to be known, for it contains $[H^+]$ (which we assume to be directly determined), $[OH^-]$ (which is determined by $[H^+]$ and the ionization constant of water), the anions of the metal salt and any strong acids and the cations of any strong bases added to the system. In practice [M], the concentration of free metal ion often can be determined experimentally and there is a method[3] based predominantly on this fact. We are interested here in the situation in which $[H^+]$, or pH, is the experimentally determined quantity.[7] Under these conditions there are $(n + \alpha + 2)$ concentrations ($[\varphi H_k]$, $k = 0, 1, \ldots, n$ and $[M\varphi_k]$, $k = 0, 1, \ldots, \alpha$) yet to be determined. If the ionization constants of φH_n are known, the corresponding mass action expressions for (2) furnish n relations and these with (3), (4) and (5) provide $(n + 3)$ relations. Now it happens that, although

(1) Presented, in part, at the 119th Meeting of the American Chemical Society at Cleveland, Ohio, Abstracts of papers, p. 39c, April, 1951.

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

(3) I. Leden, *Z. physik. Chem.*, **A188**, 160 (1941).

(4) S. Fronaeus, *Acta Chem. Scand.*, **4**, 72 (1950).

(5) J. C. Sullivan and J. C. Hindman, THIS JOURNAL, **74**, 6091 (1952).

(6) J. B. Gilbert, M. C. Otey and J. Z. Hearon, *ibid.*, **77**, 2599 (1955).

(7) It is of course assumed that b and c are known and the situation referred to here corresponds to the titration of a solution of the ampholyte and metal salt with strong base, the titration of the ampholyte with the hydroxide of M, the titration of an acid solution of M with the ampholyte, etc., with a determination of $[H^+]$ or pH after each addition.

the number of variables exceeds the number of constraints by $\alpha - 1$, a solution for $[\varphi]$ and [M] can be obtained. From (3), (4) and (5)

$$mc - pb = S + \sum_{k=0}^{n} k[\varphi H_k] \quad (6)$$

The mass action expressions for (2) can be put in the form

$$[\varphi H_k] = [\varphi][H^+]^k/P_k,\ k = 0, 1, \ldots, n \quad (7)$$

where $P_0 = 1$ (from (7)) and

$$P_k = \prod_{i=n+1}^{n+k} K_{i-k},\ 1 \leq k \leq n \quad (8)$$

and K_k is the dissociation constant for the kth equilibrium of (2). If the function $\psi(H^+)$ is defined by

$$\psi = \sum_{k=0}^{n} k[H^+]^k/P_k \quad (9)$$

then from (6), (7) and (9)

$$[\varphi] = (mc - pb - S)/\psi \quad (10)$$

Thus, when $[H^+]$ has been determined, $[\varphi]$ can be computed if the relevant equilibria are described by (1) and (2). But (10) is in fact valid, as later discussed, under much broader conditions than we have admitted for present considerations. From (7) it is clear that if the function $f(H^+)$ is defined by

$$f = \sum_{k=0}^{n} [H^+]^k/P_k \quad (11)$$

the sum of the concentrations of those φ-species which are *not* bound to M is $[\varphi]f$. The mean number of ligands bound to M is then

$$\bar{n} = (c - [\varphi]f)/b \quad (12)$$

and, as is well known, $\bar{n}$ is also given by

$$\bar{n} = \sum_{i=0}^{\alpha} i[M\varphi_i] \Big/ \sum_{i=0}^{\alpha} [M\varphi_i] = \sum_{i=0}^{\alpha} iQ_i[\varphi]^i \Big/ \sum_{i=0}^{\alpha} Q_i[\varphi]^i \quad (13)$$

where $Q_0 = 1$, $Q_i = K_1^\varphi K_2^\varphi \ldots K_i^\varphi$, $1 \leq i \leq \alpha$, and K_k^φ is the formation constant for the kth complex in (1). If $g(\varphi)$ is defined as

$$g(\varphi) = \sum_{i=0}^{\alpha} Q_i[\varphi]^i \quad (14)$$

it is evident from (13) that $\bar{n}/[\varphi] = \mathrm{d} \ln g/\mathrm{d}[\varphi]$, and thus that

$$\ln g(\varphi) = \int_0^{[\varphi]} \frac{\bar{n}}{[\varphi]}\, \mathrm{d}[\varphi] \quad (15)$$

Since

$$[M\varphi_k] = [M]Q_k[\varphi]^k,\ k = 0, 1, \ldots, \alpha \quad (16)$$

(4) may be written as

$$b = [M]g(\varphi) \quad (17)$$

and the concentration of free metal, [M], can be computed, from (10), (12), (15) and (17) when $[H^+]$ is known. The calculation of [M] under these conditions is of some interest[8] *per se*. It has been noted[5] that one limitation of the method of Leden[3] is that it assumes [M] to be experimentally measured and $[\varphi]$ to be known (or calculable by an approximate method). It has just been shown that, when the conditions assumed above apply, [M] and $[\varphi]$ are calculable essentially from $[H^+]$. For later use we note here that, from (3) and (4), with the definition (11) and (7) and (16)

$$[M] = (c - [\varphi]f - \rho b) \Big/ \sum_{k=0}^{\alpha} (k - \rho)Q_k[\varphi]^k \quad (18)$$

where ρ is any number.

III. Calculation of the Association Constants

General.—There are various alternative methods for computing the association constants when any given quantity, or set of quantities, is known from experiment. In particular it is to be noted that if the concentration of any complex, say $[M\varphi_k]$, is known the corresponding over-all constant, Q_k, can be determined at once from (16) when $[\varphi]$ and [M] have been computed from (10) and (17). In this calculation, α need not be known in advance and a value of Q_k is determined for each value of $[\varphi]$ and [M] available, *i.e.*, at each point on the titration curve.[7] Such a "point-wise" calculation is a feature of special interest in the methods to follow. When available, the knowledge of a given $[M\varphi_k]$, and the corresponding Q_k obtained as above, can be used in an obvious way to extend or modify the methods to follow.

If [M] is experimentally known, then (17) gives $g(\varphi)$ and the determination of the Q_j from this function is essentially the method of Leden.[3] Alternatively, $g(\varphi)$ can be computed from (15) and estimation of the Q_j from the function so obtained is the method of Fronaeus.[4] While these methods are now well known, we note the following. From (14)

$$(g(\varphi) - 1)/[\varphi] = Q_1 + Q_2[\varphi] + Q_3[\varphi]^2 + \ldots \quad (19)$$

and

$$(g(\varphi) - Q_1[\varphi] - 1)/[\varphi]^2 = Q_2 + Q_3[\varphi] + \ldots \quad (20)$$

Over a considerable $[\varphi]$-range, the functions (19) and (20) are frequently linear[6] and this will always be the case when there is sufficient spread among the K_j^φ. Thus Q_1 and Q_2 are estimated by the intercept and slope of the linear portion of (19), and Q_2 and Q_3 are similarly estimated from (20). It is usually true that Q_2 is better estimated by the slope of the linear portion of (19) than by the intercept of (20) and in fact intercepts of the successive functions formed from $g(\varphi)$ after the manner of (19) and (20) are quite uncertain. While undoubtedly a least-square adjustment[9] of $g(\varphi)$ is the optimum procedure, satisfactory results are obtained as outlined above[6] at least when $\alpha \leq 3$. When $\bar{n}$ can be calculated in the low $[\varphi]$-range, it is seen from (13) that

(8) *E.g.*, in the type of problem discussed by J. Z. Hearon, A. L. Schade, H. Levy and D. Burk, *Cancer Research*, **7**, 713 (1947), and by A. L. Schade, *J. Bact.*, **58**, 811 (1949).

(9) The standard assumptions of regression theory are *not* here satisfied, *i.e.*, it is not true that $[\varphi]$ is an independent variable assigned without sensible error and $g(\varphi)$ and the left-hand sides of (19) and (20) are *observed* variables subject to random error. Still such a procedure is superior to extrapolative methods previously used[3,4] and to the procedure[5] of computing the Q_j from α-values of $[\varphi]$ and (19). If the $[\varphi]$-range is not sufficient, certain Q_j will be indistinguishable (statistically) from zero (*e.g.*, ref. 6) and the number of terms in $g(\varphi)$ which are *justified by the data* can be ascertained by standard least-square theory.

$$\lim_{[\varphi]\to 0}\left\{\frac{\bar{n}}{[\varphi]}\right\} = Q_1 \qquad (21)$$

and

$$\lim_{[\varphi]\to 0}\left\{\frac{\partial\bar{n}/[\varphi]}{\partial[\varphi]}\right\} = 2Q_2 - Q_1^2 \qquad (22)$$

furnish graphical estimates of Q_1 and Q_2.

The System $\alpha = 1$.—Although this is not a frequent case, it is of interest to note that *only* in this case the $(n + 3)$ relations, (3), (4), (5) and (7), suffice to determine the relevant concentrations ($[\varphi H_k]$, $k = 0, 1, \ldots, n$, and $[M\varphi_k]$, $k = 0, 1$) which are $n + 3$ in number. From (3) and (7), with the definition (11)

$$[M\varphi] = c - [\varphi]f \qquad (23)$$

while (18), with $\rho = 1$, $\alpha = 1$, gives

$$[M] = b - c + [\varphi]f \qquad (24)$$

Thus the relation

$$Q_1 = (c - [\varphi]f)/(b - c + [\varphi]f)[\varphi] \qquad (25)$$

with $[\varphi]$ from (10), provides a point-wise calculation of Q_1. Equation (25), with (10), is a generalization of special equations previously presented for this case.[10]

The System $\alpha = 2$.—From (18), it is seen that, for any positive value of ρ, the numerator of that expression may vanish for a certain value of $[\varphi]$, *viz.*, $[\varphi] = (c - \rho b)/f$, and this is clearly possible whenever $c > \rho b$. Since [M] is not zero for any finite $[\varphi]$, the denominator of (18) must vanish at this same value of $[\varphi]$. This fact (discussed in detail elsewhere[11] for a special case) imposes a relation among certain of the Q_j. We consider here only the case $\rho = 1$ and the relation referred to is then

$$\sum_{k=0}^{\alpha}(k - 1)Q_k[\varphi]_0^k = 0 \qquad (26)$$

which we call, for obvious reasons, the indeterminate point relation, where $[\varphi]_0$ is determined (in conjunction with (10)) by

$$[\varphi]_0 = (c - b)/f \qquad (27)$$

In particular, when $\alpha = 2$, (26) determines Q_2 as

$$Q_2 = 1/[\varphi]_0^2 \qquad (28)$$

This fact has been previously noted[11] and employed as a method[11,12] of estimating Q_2. For $\alpha = 2$, then, Q_1 and Q_2 can be determined from (21) and (28). This method gives a single estimate of Q_1 and Q_2 for each titration curve, but the extrapolation (21) is quite reliable and $[\varphi]_0$ can be accurately determined by computing $[\varphi]$, by (10), for points on the smoothed titration curve and locating the point at which (27) holds. However, using the relation so far derived, methods are available which permit a point-wise determination of Q_1, Q_2 or both.

Case I.—In this case, Q_2 is determined from (28) and a point-wise calculation made for Q_1. With Q_2 known, (18) determines [M], at any $[\varphi] \neq [\varphi]_0$, as

$$[M] = (b - c - [\varphi]f)/(1 - Q_2[\varphi]^2) \qquad (29)$$

From (4)

$$[M\varphi] = b - [M](1 + Q_2[\varphi]^2) \qquad (30)$$

For each value of [M] and $[M\varphi]$ computed from (29) and (30), Q_2 can be computed from (16). Plainly the values of [M], from (29) are uncertain in the immediate neighborhood of the indeterminate point, $[\varphi] = [\varphi]_0$ but $[\varphi]_0$, and hence Q_2, can be determined in several experiments and an average value of Q_2 so determined can be used, in (29) and (30), with data from experiments in which the indeterminate point does *not* occur ($b \geq c$), and values from (29) are reliable over the entire concentration range.

Case II.—In this case, Q_1 is determined by (21) and a point-wise calculation is made for Q_2. We define, for use here and elsewhere, the set of ratios

$$R_k = [M\varphi_k]/[M] = Q_k[\varphi]^k,\ k = 0, 1, \ldots \alpha \qquad (31)$$

For each $[\varphi]$ for which the R_k are determined, the Q_k can be computed. With Q_1 (and hence R_1) known, (3) and (4) furnish two equations

$$R_1[M] + 2[M\varphi_2] = \bar{n}b \qquad (32)$$

$$(1 + R_1)[M] + [M\varphi_2] = b \qquad (33)$$

which determine R_2 as

$$R_2 = \{R_1 - \bar{n}(1 + R_1)\}/(\bar{n} - 2) \qquad (34)$$

Case III.—In this case, $g(\varphi)$ is determined from (15). This by (17) is tantamount to knowing [M] and (3) and (4) provide two relations determining $[M\varphi]$ and $[M\varphi_2]$. The numerical evaluation of (15) is somewhat more laborious than the use of (21) or (28), but the subsequent numerical work is extremely simple and the method furnishes point-wise calculations of Q_1 and Q_2. With (17) and the definition (31), (32) and (33) become

$$R_1 + 2R_2 = \bar{n}g \qquad (35)$$

$$R_1 + R_2 = g - 1 \qquad (36)$$

and the solutions are

$$R_1 = (2 - \bar{n})g - 2 \qquad (37)$$

$$R_2 = 1 + (\bar{n} - 1)g \qquad (38)$$

For each $[\varphi]$ computed from (10), $\bar{n}$ is evaluated from (12), $g(\varphi)$ from (15), R_1 and R_2 from (37) and (38), and Q_1 and Q_2 from (31).

The System $\alpha = 3$.—For this system, with $[\varphi]$ determined by (10), we seek two relations which, with (3) and (4), serve to determine the concentration $[M\varphi_k]$, $k = 0, 1, 2, 3$. There are three main alternatives.

Case I.—In this case, Q_1 is determined from (21) and the remaining necessary relation is taken as (26) with $\alpha = 3$. For use here and elsewhere, we define the ratios

$$\gamma_k = [\varphi]_0^k/[\varphi]^k,\ k = 0, 1, \ldots, \alpha \qquad (39)$$

Then (3), (4) and (26) become

$$R_1[M] + 2[M\varphi_2] + 3[M\varphi_3] = \bar{n}b \qquad (40)$$

$$(1 + R_1)[M] + [M\varphi_2] + [M\varphi_3] = b \qquad (41)$$

$$[M] - \gamma_2[M\varphi_2] - 2\gamma_3[M\varphi_3] = 0 \qquad (42)$$

which may be solved for $[M\varphi_k]$, $k = 0, 2, 3$; but the ratios R_2 and R_3 can be written down at once as the ratios of the appropriate determinants. They are

(10) For example eq. 10 of S. Chaberek, Jr., and A. E. Martell, THIS JOURNAL, **74**, 6228 (1952).

(11) J. Z. Hearon, D. Burk and A. L. Schade, *J. Nat. Cancer Inst.*, **9**, 337 (1949).

(12) A. Albert, *Biochem. J.*, **47**, 531 (1950).

$$R_2 = \frac{2\gamma_3[\bar{n} + R_1(\bar{n} - 1)] + (\bar{n} - 3)}{(\bar{n} - 2)(\gamma_2 - 2\gamma_3) - \gamma_2} \quad (43)$$

$$R_3 = \frac{\gamma_2[R_1(1 - \bar{n}) - \bar{n}] + (2 - \bar{n})}{(\bar{n} - 2)(\gamma_2 - 2\gamma_3) - \gamma_2} \quad (44)$$

For each $[\varphi]$ computed from (10), $\bar{n}$ is evaluated from (12), the γ_i from (39), R_2 and R_3 from (43) and (44), and Q_2 and Q_3 from (31).

Case II.—In this case, Q_1 is determined from (21) and the remaining necessary relation is taken as (17), with $g(\varphi)$ from (15). Then (40) and (41) can be written as

$$2R_2 + 3R_3 = \bar{n}g - R_1 \quad (45)$$

$$R_2 + R_3 = g - 1 - R_1 \quad (46)$$

The solutions are

$$R_2 = (3 - \bar{n})g - (3 + 2R_1) \quad (47)$$

$$R_3 = (\bar{n} - 2)g + 2 + R_1 \quad (48)$$

For each $[\varphi]$, $\bar{n}$ and $g(\varphi)$ are computed from (12) and (15), R_2 and R_3 from (47) and (48), and Q_2 and Q_3 from (31).

Case III.—In this case (26), with $\alpha = 3$, and (17), with $g(\varphi)$ from (15), are used in conjunction with (3) and (4). With the definitions (31) and (39), (26) becomes

$$\gamma_2 R_2 + 2\gamma_3 R_3 = 1 \quad (49)$$

and (49), (45) and (46) determine R_k, $k = 1, 2, 3$. From (45), (46) and (49)

$$R_1 = \frac{(2\gamma_3 - \gamma_2)\bar{n}g + (g - 1)(3\gamma_2 - 4\gamma_3) - 1}{2(\gamma_2 - \gamma_3)} \quad (50)$$

With R_1 computed from (50), R_2 and R_3 are given by

$$R_2 = \{(g - 1 - R_1)2\gamma_3 - 1\}/(2\gamma_3 - \gamma_2) \quad (51)$$

$$R_3 = \{1 - \gamma_2(g - 1 - R_1)\}/(2\gamma_3 - \gamma_2) \quad (52)$$

or alternatively by substituting R_1 from (50) into (47) and (48). For each $[\varphi]$, $\bar{n}$ is computed from (12), g from (15), the γ_k from (39), the R_k from (50)–(52), [or from (50), (47) and (48)], and a point-wise calculation of the Q_k is made from (31).

The Systems $\alpha \geq 4$.—If (3), (4) and the equation corresponding to (26), but with general ρ, are written as

$$\left.\begin{aligned} \sum_{k=0}^{\alpha} kR_k &= \bar{n}g \\ \sum_{k=0}^{\alpha} R_k &= g \\ \sum_{k=0}^{\alpha} (k - \rho)\gamma_k R_k &= 0 \end{aligned}\right\} \quad (53)$$

where the γ_k are defined by (39) and $[\varphi]_0$ is now given by $(c - \rho b)/f$, it is apparent that, if g is known from (15), additional relations, $(\alpha - 3)$ in number, are required to determine the R_k. If Q_1 is determined from (21), R_1 is known and (53) suffices to determine the remaining Q_j, for each $[\varphi]$, when $\alpha = 4$. If Q_1 and Q_2 are determined from (21) and (22), R_1 and R_2 are known and (53) determines, for each $[\varphi]$, the remaining Q_j when $\alpha = 5$. If $\alpha > 5$, Q_1 and Q_2 may be determined from (21) and (22) then (53), with two or more different values of ρ, permits a pointwise calculation of the Q_k, $k = 3, 4, 5, 6$. In general many such combinations can be employed, *e.g.*, (21) and/or (22) may be dispensed with and two or three or more values of ρ assigned in (53), etc. It may be noted that when there is a sufficient spread among the K_j^{φ}, (26), written as (28), gives a good *approximation* to Q_2 (*cf.* discussion in ref. 6).

IV. Calculation of $[\varphi]$ from the Association Constants

When estimates of the Q_j have been obtained, it is frequently of interest to compute the values of $[\varphi]$, [M] and $[M\varphi_k]$, $k = 1, 2, \ldots, \alpha$, which obtain in a system for which b, c and $[H^+]$, or pH, are prescribed. This represents one practical value of the known Q_j. Further, a comparison of the $[\varphi]$-values so computed with those obtained from the experimental data indicates how well the estimated Q_j reproduce the experimental data. In a system at equilibrium, the prevailing value of $[\varphi]$ is determined from (12) and (13) by

$$(c - [\varphi]f)/b = \sum_{i=0}^{\alpha} iQ_i[\varphi]^i \Big/ \sum_{i=0}^{\alpha} Q_i[\varphi]^i \quad (54)$$

which in general is of degree $(\alpha + 1)$ in $[\varphi]$ and cannot be solved by (non-iterative) algebraic methods. But for fixed b, c and $[H^+]$, the solution of (54) is clearly that value of $[\varphi]$ for which the line

$$y = (c - [\varphi]f)/b \quad (55)$$

sects the curve represented by the right-hand side of (54). From the known Q_j the curve can be plotted and the line (55) is easily constructed by connecting the ordinate intercept, c/b, and the intercept, c/f, on the $[\varphi]$-axis.[13] The values of $[\varphi]$ so obtained can be compared to those computed from (10). In constructing the curve, given by the right-hand side of (54), the function $g(\varphi)$ must be computed, and from these values [M] is given by (17). The concentrations $[M\varphi_k]$ can now be computed from (16).

A quantity of some interest[6] is the number of M bound per mole of ampholyte. This quantity, n^*, is

$$n^* = (b - [M])/c \quad (56)$$

which with (17) can be written as

$$n^* = b(g - 1)/cg \quad (57)$$

When $[\varphi]$ and $g(\varphi)$ have been computed as above, n^* is readily calculated.

Under certain conditions, relatively simple expressions for the ratio of two analogous complexes $[M\varphi_k^a]/[M\varphi_k^b]$ can be obtained, where φ^a and φ^b are the completely ionized forms of two ampholytes. From (16)

$$[M\varphi_k^a]/[M\varphi_k^b] = Q_k^a[\varphi^a]^k/Q_k^b[\varphi^b]^k \quad (58)$$

where Q_k^a and Q_k^b are the over-all association constants for the type-a and type-b ampholytes. If the total concentrations of the ampholytes are c_a and c_b and both of these quantities are very large relative to b, then it follows from (3) and (4) that

$$[\varphi^a] = c_a/f_a \quad (59)$$

where f_a is given by (11) for the type-a ampholyte and a similar expression holds for $[\varphi^b]$.

(13) In the most common experimental situation b and c are fixed and f takes a given value for each $[H^+]$. The necessary system of lines is constructed by connecting the fixed ordinate-intercept, c/b, with the abscissas c/f, there being one such point for each $[H^+]$ on the titration curve (examples are given in ref. 6). The ordinates of the points of intersection are the computed values of $\bar{n}$.

V. Discussion

Equation 10 represents the solution of (3), (4) and (5) for $[\varphi]$ for any α, n, m and p, where in particular m may be zero (*e.g.*, when φ is NH_3, ethylenediamine, etc.). Many investigators (to cite a few, 10, 11, 12, 14–18), of whom the majority used the method of Bjerrum,[2] have correctly computed $[\varphi]$ from the two material balances and the electroneutrality condition[19] and the equations which they have presented are in each case special versions of (10), the forms of which depend upon the valences and special experimental conditions.

The methods suggested here permit a pointwise calculation of certain, in some cases all, of the Q_j and thus reveal any trend in, or lack of constancy of, these quantities. When suitably constant, over the entire concentration range, the Q_j are obtained as the average of as many *independent* determinations as there are points on the titration curve. As in any mass-action computation, there are regions of relative uncertainty when the relevant quantities are small differences between large numbers. The formulations here and the discussion of the indeterminate point reveal clearly such regions and why they occur. From (12), (13) and (18) it is plain that the indeterminate point occurs when $\bar{n} = \rho$ and there is perhaps some merit in assigning $\rho \geq 1$ when $\alpha \geq 3$, or in employing several values of ρ as suggested in the last section. The method[5,20] of computing the Q_j, using (13) in the form $\sum_0^{\alpha}(k - \bar{n})Q_k [\varphi]^k = 0$, from α values of $[\varphi]$ is likewise plagued by these regions of uncertainty and gives no indication of trend, in general, unless α is small and many data points are available.[21]

There are surely uncertainties in the estimation of slopes, intercepts or, by interpolation, particular values of $\bar{n}$ or $[\varphi]$. But it is worth noting that (in common with the relations suggested by Scatchard[18] and in contrast to those suggested by Bjerrum[2]) relations (20), (21) and (26) or (28) are mathematically exact. The Q_j determined by these relations do not *necessarily* require subsequent adjustment.

Finally we point out that (10) and (54) enable the presence of polynuclear and proton-bearing complexes to be ruled out in a given system. This appears to be important for such complexes do occur.[22] The right-hand side of (54) is a function of $[\varphi]$ alone (*i.e.*, $\bar{n}$ is fixed when $[\varphi]$ is prescribed regardless of what conditions of total concentrations and pH produced that value of $[\varphi]$) and the point at which that curve is intersected by the line (55) is determined by the three quantities b, c and f. If these three quantities are changed in such a way that c/b and c/f remain fixed, then plainly the value of $[\varphi]$ is unchanged, for the slope and intercept of (55) are unchanged. This property is strictly a consequence of the assumption, embodied in (13), that polynuclear and proton-bearing complexes are not formed, for otherwise the right-hand side of (54) would be an *explicit* function of $[H^+]$ and b. Consider a system in which a metal ion–ampholyte mixture has been adjusted in pH until b_1, c_1 and f_1 are the prevailing values of the total concentrations and the function (11). Now let more metal salt and ampholyte be added and the pH adjusted so that the prevailing values are b_2, c_2 and f_2 where $c_1/b_1 = c_2/b_2$ and $c_1/f_1 = c_2/f_2$. According to the foregoing the value $[\varphi]_1$ computed from (10) with b_1, c_1 and ψ_1 should agree with the value $[\varphi]_2$ computed from (10) with b_2 and c_2 and ψ_2. Alternatively b_2, c_2 and ψ_2 may be substituted into (10), the result equated to $[\varphi]_1$, and the titration value (contained in S) computed. This *predicts* the titration value at which the pH is such that $c_1/f_1 = c_2/f_2$ and this can be checked against the observed value. There are many ways in which the test can actually be carried out and the range of pH and metal concentration over which the system is described by (1) and (2) can be established. It should be noted that (10) is correct even when *certain* types of polynuclear complexes are formed but never when any proton-bearing complexes are formed. Denote by X_{kj} the complex $\varphi_k M_j G$ where G is any ion of valence zero or group of ions whose valences sum to zero. Then from (3), (4) and (5), written as

$$c = [\varphi]f + \Sigma\Sigma k X_{kj}$$
$$b = \Sigma\Sigma j X_{kj}$$
$$S + [\varphi]\psi = \Sigma\Sigma(mk - pj)X_{kj} + m[\varphi]f$$

where the summations run over all possible values but $j \geq 1$, it is evident at once that (10) results. Thus (10) is correct when the valences of the complexes formed are determined only by the number of φ's and M's which the complexes contain. Since

(14) H. Flood and V. Loras, *Tids. Kjemi, Bergvesen Met.*, **5**, 89 (1945).

(15) M. Calvin and M. K. Wilson, THIS JOURNAL, **67**, 2003 (1945).

(16) H. B. Jonassen and A. W. Meibohm, *J. Phys. Colloid Chem.*, **55**, 726 (1951).

(17) S. Chaberek, Jr., and A. E. Martell, THIS JOURNAL, **74**, 5052 (1952).

(18) J. T. Edsall, G. Felsenfeld, D. S. Goodman and F. R. N. Gurd, *ibid.*, **76**, 3054 (1954).

(19) It is not clear that Bjerrum,[2] in his treatment of systems in which equilibria (2) are important, invokes electroneutrality. Bjerrum[2] proceeds (p. 202) through the unnecessary device of defining C'_{en}, C_s, α_{en} and $\bar{n}_{en}$, which generalized in our terminology are $[\varphi]f$, $[\varphi]\psi$, $1/f$ and ψ/f, respectively. His calculation of $\bar{n}$ (eq. 6, p. 203) pivots upon the relation $C_s/\alpha_{en} = [\varphi]f$, which from our point of view is redundant but not incorrect. However, C_s (termed the "concentration of acid bound") is stated (p. 203) to be the "total acid concentration minus the concentration of free hydrogen ions." This assertion is, strictly speaking, incompatible with electroneutrality as is easily verified from our (10) and reference to Bjerrum's experimental conditions. In fact, the correct version of Bjerrum's C_s has since appeared in the literature.[16] Although for *practical* purposes Bjerrum's statement is correct, due to the coincidence that in his case $m = 0$ and $[H^+]$ and $[OH^+]$ are negligible relative to the concentration of strong acid added to the system, this is not of necessity true and it seems important to realize the logical status of his formulation.

(20) B. P. Block and G. H. McIntyre, THIS JOURNAL, **75**, 5667 (1953).

(21) If a group of points, α in number, are used to compute the Q's then a given Q_j is not obtained over a wide $[\varphi]$-range unless successive groups of points overlap, in which case the determinations are not independent, or the $[\varphi]$-values are closely spaced. In the latter case, the determinant of the system is "nearly singular" and unreliable values are obtained. We are aware of the fact that mathematically speaking the determinant is non-singular under any sensible experimental conditions (*cf.* 5 and 20) but for practical numerical purposes the determinant may be "nearly singular." This is well known and such systems, called "ill conditioned," have been the subject of much concern and research (*e.g.*, Q. J. Turing, *Quart. J. Mech. Appl. Math.*, **1**, 287 (1948)). The work reported here was elicited by precisely these difficulties in computing the Q_j from a least-square fit of eq. (13) in the form mentioned above or from α simultaneous equations.

(22) R. C. Warner and I. Weber, THIS JOURNAL, **75**, 5086 (1953). Also discussion in refs. 5 and 11.

(12) is *always* correct, $[\varphi]$ and $\bar{n}$ can always be computed for such systems. Certain examples of such systems[11] involve oxygenated complexes (*e.g.* $(M\varphi_2)_2O_2$) and have been thoroughly studied.

NOTE ADDED IN PROOF.—It is clear enough that we have excluded from consideration throughout this paper *all* species not involved in equilibria (1) and (2). Polynuclear and proton-bearing complexes were explicitly excluded because they are common and because their presence can be ruled out by tests based on the present formulation. However, in view of a recent report (C. Tanford, D. C. Kirk and M. K. Chantooni, THIS JOURNAL, **76**, 5325 (1954)) hydroxyl ion-bearing complexes, of the general form $M\varphi_k(OH)_j$ with valence $p - km - j$, are to be explicitly excluded. The presence of such complexes also can be ruled out on the basis of the tests referred to above. Obviously these tests preclude the existence of no complex but permit a decision as to whether certain complexes are present in quantitatively important amounts.

OAK RIDGE, TENNESSEE
BETHESDA 14, MARYLAND

[CONTRIBUTION FROM THE LABORATORY OF BIOCHEMISTRY, NATIONAL CANCER INSTITUTE, NATIONAL INSTITUTES OF HEALTH, PUBLIC HEALTH SERVICE AND THE DEPARTMENT OF PHYSIOLOGY, UNIVERSITY OF CHICAGO]

Association Constants of Cobalt–Glycine and Cobalt–Glycylglycine Complexes in Aqueous Solution[1,2,3]

BY JAMES B. GILBERT, M. CLYDE OTEY AND JOHN Z. HEARON

RECEIVED AUGUST 2, 1954

Association constants for the complexes of cobaltous ion with glycine, glycylglycine and glycyl-D-alanine have been determined. The new methods of computation[8] of complex ion equilibria have been applied to potentiometric titration data with consistent and reliable results. The relative quantities of cation and anionic chelating agent bound in complex ion forms have been computed. The results indicate that glycine has greater affinity for cobaltous ion than does glycylglycine. This finding is contrary to evidence which has been employed to support a hypothesis of enzyme–cobalt–glycylglycine complex formation. This complex had been postulated as the intermediate explaining the metal ion activation of the enzymatic cleavage of glycylglycine.

Introduction

A possible explanation of the cobaltous ion activation of the enzymatic hydrolysis of glycylglycine has been advanced. This theory has postulated an enzyme–cobalt–glycylglycine complex as an integral part of the proposed mechanism.[4] This suggestion was made on the basis of evidence that indicated a strong degree of interaction of cobaltous ion with glycylglycine, but not with glycine, glycylglycylglycine or numerous other peptides.[5]

The evidence proposed consisted of the demonstration of strong red coloration of solutions containing glycyglycine and cobaltous ion when allowed to stand at *p*H 8 under aerobic conditions. However, this has been shown to be indicative of the capacity of previously formed complexes to oxygenate.[6,7] Furthermore the red oxygenated complexes of the peptides examined including glycylglycine are not hydrolyzed enzymatically. Therefore, it is of interest, particularly with respect to the suggested theory of enzymatic activation, to compare under anaerobic conditions, the relative affinities of glycylglycine, the substrate, and glycine, the end-product, for cobaltous ion.

This problem also supplies suitable chelating anion and cation to test the validity of the new methods[8] for computation of chelation association constants.

Theoretical.—Equations employed have been described.[8] However, it is noted that equation 10 simplifies to

$$[\phi^-] = \frac{(c - t) + K_w/[H^+] - [H^+]}{[H^+]/K_2 + [H^+]^2/K_1K_2} \qquad (101)^8$$

where t is the concentration of sodium ion in solution, $[H^+]$ is the hydrogen ion activity[9] as measured by the glass electrode, and K_w is the ion product for water.[10] The evaluation of Q_2 from equation 28 is accomplished by doing a "point-by-point" calculation along a given titration curve until the numerator of the right member of equation 18 vanishes, *i.e.*, equation 27 is satisfied.

Experimental

Materials.—Glycylglycine was synthesized by amination of chloroacetylglycine.[11] *Anal.* Calcd. for $C_4H_8N_2O_3$: N, 21.2. Found: N, 21.2. Glycyl-D-alanine was prepared from chloroacetyl-D-alanine obtained by enzymatic resolution.[12] *Anal.* Calcd. for $C_5H_{10}N_2O_3$: N, 19.2. Found: N, 19.0. Glycine was twice recrystallized from water. *Anal.* Calcd. for $C_2H_5NO_2$: N, 18.7. Found: N, 18.8. $CoCl_2 \cdot 6H_2O$, analytical reagent grade, was used throughout.

Apparatus and Methods.—A Beckman model G *p*H meter with extension glass and calomel electrodes was employed to measure *p*H. Anaerobic conditions were maintained in the titration vessel and also in the sodium hydroxide (carbonate-free) by circulating nitrogen previously passed over

(1) The subject matter of this paper was presented, in part, at the 119th National Meeting of the American Chemical Society at Cleveland, Ohio, April, 1951.

(2) Abstracts of Meeting Papers, p. 39c, April, 1951.

(3) This work was aided, in part, by a grant to (J.Z.H.) from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of the University of Chicago.

(4) E. L. Smith, *J. Biol. Chem.*, **176**, 21 (1948).

(5) E. L. Smith, *ibid.*, **173**, 571 (1948).

(6) J. B. Gilbert, M. C. Otey and V. E. Price, *ibid.*, **190**, 377 (1951).

(7) J. Z. Hearon, D. Burk and A. L. Schade, *J. Nat. Cancer Inst.*, **9**, 337 (1949).

(8) J. Z. Hearon and J. B. Gilbert, THIS JOURNAL, **77**, 2594 (1955). References to equations numbered with one or two digits indicate equations of this preceding paper. Three digit equation references are contained in the present paper.

(9) Note that $[H^+]$ is in terms of activity throughout this paper while other factors are in terms of concentration on a molar basis.

(10) Note that most terms have been previously defined. However, the appropriate simplifications are taken, in this paper, *e. g.*, $[\phi^{-m}] = [\phi^-]$, $[M^{+p}] = [Co^{++}]$, without further explanation.

(11) E. Fischer, *Ber.*, **37**, 2486 (1904).

(12) P. J. Fodor, V. E. Price and J. P. Greenstein, *J. Biol. Chem.*, **178**, 503 (1949).