

- Cohn, M., and Urey, H. C. (1938), *J. Am. Chem. Soc.* 60, 679.
- Ellman, G. L. (1959), *Arch. Biochem. Biophys.* 82, 70.
- Evans, R. L., and Saroff, H. A. (1957), *J. Biol. Chem.* 228, 295.
- Feeney, R. E., Osuga, D. T., and Maeda, H. (1967), *Arch. Biochem. Biophys.* 119, 124.
- Fischer, E. H., Kent, A. B., Snyder, E. R., and Krebs, E. G. (1958), *J. Am. Chem. Soc.* 80, 2906.
- Fraenkel-Conrat, H. (1957), *Methods Enzymol.* 4, 247.
- Fraenkel-Conrat, H., and Olcott, H. S. (1948), *J. Biol. Chem.* 174, 827.
- Gorecki, M., and Shalitin, Y. (1967), *Biochem. Biophys. Res. Commun.* 29, 189.
- Habeeb, A. F. S. A. (1966), *Anal. Biochem.* 14, 328.
- Haynes, R., Osuga, D. T., and Feeney, R. E. (1967), *Biochemistry* 6, 541.
- Hirs, C. H. W. (1962), *Brookhaven Symp. Biol.* 15, 154.
- Hughes, W. L., Saroff, H. A., and Carney, A. L. (1949), *J. Am. Chem. Soc.* 71, 2476.
- Hunter, M. J., and Ludwig, M. L. (1962), *J. Am. Chem. Soc.* 84, 3491.
- Ingram, V. M. (1953), *J. Biol. Chem.* 202, 193.
- Kallen, R. G., and Jencks, W. P. (1966), *J. Biol. Chem.* 241, 5864.
- Klee, W. A., and Richards, F. M. (1957), *J. Biol. Chem.* 229, 489.
- Kunitz, M. (1946), *J. Biol. Chem.* 164, 563.
- Levy, A. L., and Li, C. H. (1955), *J. Biol. Chem.* 213, 487.
- Marfey, P. S., Nowak, H., Uziel, M., and Yphantis, D. A. (1965), *J. Biol. Chem.* 240, 3264.
- Martin, C. J., and Marini, M. A. (1967), *J. Biol. Chem.* 242, 5736.
- Moffit, W., and Yang, J. T. (1956), *Proc. Natl. Acad. Sci. U. S.* 42, 596.
- Moore, S., Cole, R. D., Gundlach, H. G., and Stein, W. H. (1958), *4th Intern. Congr. Biochem.*, 52.
- Moore, S., and Stein, W. H. (1948), *J. Biol. Chem.* 176, 367.
- Okuyama, T., and Satake, K. (1960), *J. Biochem. (Tokyo)* 47, 454.
- Paik, W. K., and Kim, S. (1967), *Biochem. Biophys. Res. Commun.* 27, 479.
- Schütte, E. (1943), *Z. Physiol. Chem.* 279, 59.
- Shields, G. S., Hill, R. L., and Smith, E. L. (1959), *J. Biol. Chem.* 234, 1747.
- Stevens, F. C., and Feeney, R. E. (1963), *Biochemistry* 2, 1346.
- Tanford, C., and Hauenstein, J. D. (1956), *J. Am. Chem. Soc.* 78, 5287.
- Yang, J. T., and Doty, P. (1957), *J. Am. Chem. Soc.* 79, 761.
- Zittle, C. A. (1948), *J. Franklin Inst.* 246, 266.

The Determination of Association Constants for a Two-Solute System. I. Colligative Methods*

Robert F. Steiner

ABSTRACT: A general method is outlined for obtaining association constants for a system of two different associating species, which can associate with themselves

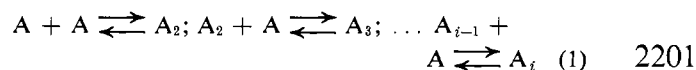
or with each other in any proportion. The theory applies to colligative methods, which yield number-average molecular weights.

The problem of determining the consecutive equilibrium constants for an associating system from molecular weight data is a recurrent one in biochemistry. While proteins furnish many examples of associating systems (Reithel, 1963), there are also numerous cases of small molecules or ions which associate in solution (Ts'o *et al.*, 1963).

Several experimental approaches are available for obtaining molecular weights for systems of this kind. Of

these, osmotic pressure and the other colligative methods yield number-average molecular weights, while light-scattering and sedimentation equilibrium yield weight-average quantities. The colligative methods other than osmotic pressure, such as freezing point depression and vapor pressure lowering, are primarily applicable to small molecules of molecular weight less than 10^3 , while the other techniques apply to biopolymers.

Several procedures are available for computing the association constants of a self-associating system, in which a single component forms dimers, trimers, etc. For an association process of the type



* From the Laboratory of Physical Biochemistry, Naval Medical Research Institute, Bethesda, Maryland. Research Task MR005.06-0005. Received February 9, 1968. The opinions in this paper are those of the author and do not necessarily reflect the views of the Navy Department or the naval service at large.

the molar concentration of each species is given in terms of a set of association constants K_i by an expression of the form

$$[A_i] = K_i[A]^i \quad (K_1 = 1) \quad (2)$$

where $[A]$ = molar concentration of free A units. The existing treatments for obtaining the values of K_i for the consecutive association steps all rely upon similar procedures for computing the mole fraction of monomer units from the observed concentration dependence of number-average or weight-average molecular weight (Steiner, 1952, 1954; Adams, 1964, 1965; Adams and Fujita, 1963; Adams and Williams, 1964). In the former case, for ideal conditions (Steiner, 1954)

$$\ln x_A = \int_0^m (\alpha_n^{-1} - 1) \frac{dm}{m} \quad (3)$$

where $x_A = [A]/m$ = mole fraction of monomer units, m = total molar concentration of all species = $\sum K_i[A]^i$, α_n = number-average degree of association = M_n/M_A , M_n = number-average molecular weight, and M_A = molecular weight of a monomer. If the total weight concentration is c , then

$$m = c/M_n \quad (4)$$

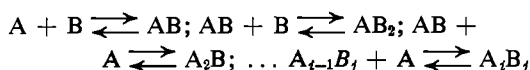
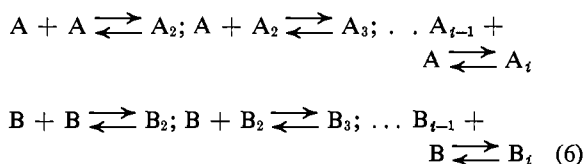
The consecutive association constants may be found from the equation (Steiner, 1954)

$$\frac{m}{[A]} = 1 + K_2[A] + K_3[A]^2 + \dots \quad (5)$$

This may be done either by graphical analysis of limiting slopes or by polynomial fitting.

A general method for the analysis of a system containing two different monomer species is not at present available. Examples of such systems are numerous in biochemistry, including the antigen-antibody reaction, the interaction of trypsin and chymotrypsin with protein inhibitors, the hemoglobin-haptoglobin interaction, the association of nonidentical subunits of an enzyme, the complexing of lysozyme with plasma albumin or ovalbumin, and the association of purines and pyrimidines in solution.

The present paper will outline one possible approach to this problem. The most general model is that of a system consisting of two molecular units, A and B, each of which may associate with itself, or with the other monomer unit, in any proportion.



Most actual cases will represent simplifications of the general model. Thus one, or both, of the monomer units may often not undergo self-association. The theory developed here will be applicable to any experimental procedure yielding number-average molecular weights.

Theory

The treatment to be presented here applies to an associating system containing two different monomeric species, which exhibits ideal behavior, so that the apparent number-average molecular weight is equivalent to the actual value.

In practice, both the total concentration and the ratio of total B to total A may be varied. In this way a series of curves of number-average molecular weight *vs.* concentration may be obtained, corresponding to a set of different over-all compositions.

The molar concentration of each species in the system described by eq 6 is given by

$$[A_iB_j] = K_{ij}[A]^i[B]^j \quad (K_{10} = K_{01} = 1; K_{00} = 0) \quad (7)$$

We then have for the total molar concentration of all species

$$m = \sum_i \sum_j K_{ij}[A]^i[B]^j \quad (8)$$

The *total* molar concentration of all monomer species is given by

$$\begin{aligned} m_t = m_{A,t} + m_{B,t} &= \sum_i iK_{i0}[A]^i + \sum_j jK_{0j}[A]^0[B]^j \\ &= [A] \frac{\partial m}{\partial [A]} + [B] \frac{\partial m}{\partial [B]} \end{aligned} \quad (9)$$

and

$$[A] \frac{\partial m}{\partial [A]} = m_{A,t}$$

$$[B] \frac{\partial m}{\partial [B]} = m_{B,t}$$

Here $m_{A,t}$ and $m_{B,t}$ are the *total* molar concentrations of A and B units, respectively, whether free or complexed.

Since c = total weight concentration = $M_n m$ and M_0 = number-average molecular weight when association is absent

$$\frac{M_A m_{A,t} + M_B m_{B,t}}{m_{A,t} + m_{B,t}} = \frac{c}{m_t} \quad (10)$$

we have

$$m_t = \frac{c}{M_0} = \alpha_n m \quad (11)$$

where $\alpha_n = M_n/M_0$ = number-average degree of association.

Substituting in eq 9

$$m\alpha_n - [A] \frac{\partial m}{\partial [A]} - [B] \frac{\partial m}{\partial [B]} = 0 \quad (12)$$

Let us for the present assume that the ratio of $m_{B,t}$ to $m_{A,t}$ is constant and equal to β . This corresponds to the experimental situation where only the total weight concentration is varied, the ratio of the total molar (or weight) concentrations of B and A remaining constant.

Then

$$m_{B,t} = \beta m_{A,t} \quad (13)$$

where β is constant and

$$[B] \frac{\partial m}{\partial [B]} = \beta [A] \frac{\partial m}{\partial [A]}$$

Substituting in eq 12

$$m\alpha_n - (1 + \beta)[A] \frac{\partial m}{\partial [A]} = 0 \quad (14)$$

or

$$m\alpha_n - (1 + \beta) \frac{\partial m}{\partial \ln [A]} = 0$$

This may be integrated to yield

$$\ln [A] = (1 + \beta) \int \frac{dm}{\alpha_n m} + F([B]) \quad (15)$$

where $F([B])$ is an arbitrary function of $[B]$. This may be placed in the form

$$G(m, [A], [B]) = 0 \quad (16)$$

where

$$G = \ln [A] - (1 + \beta) \int \frac{dm}{\alpha_n m} - F([B])$$

We then have

$$\frac{\partial m}{\partial [A]} = - \frac{\partial G / \partial [A]}{\partial G / \partial m} = \frac{\alpha_n m}{[A](1 + \beta)} \quad (17)$$

$$\frac{\partial m}{\partial [B]} = - \frac{\partial G / \partial [B]}{\partial G / \partial m} = - \frac{\alpha_n m}{1 + \beta} \frac{dF}{d[B]}$$

Substituting in eq 12

$$\alpha_n m - \frac{\alpha_n m}{1 + \beta} + \frac{[B]}{1 + \beta} \frac{dF}{d[B]} = 0 \quad (18)$$

and $F([B]) = -\beta \ln [B] + C$, where C is a constant. Thus, we have from eq 15 and 18

$$\ln [A] + \beta \ln [B] = (1 + \beta) \int \frac{dm}{\alpha_n m} + C \quad (19)$$

Since

$$[A] = x_A m; [B] = x_B m \quad (20)$$

where x_A and x_B are the mole fractions of monomeric A and B, respectively, we have

$$\ln [A] = \ln x_A + \int \frac{dm}{m} \quad (21)$$

$$\ln [B] = \ln x_B + \int \frac{dm}{m}$$

and

$$\begin{aligned} \ln x_A + \beta \ln x_B = \\ (1 + \beta) \int_0^m (\alpha_n^{-1} - 1) \frac{dm}{m} + C \end{aligned} \quad (22)$$

If association is absent, $\alpha_n = 1$ and the integral vanishes. In this case

$$x_A = \frac{m_{A,t}}{m_t} = \frac{1}{1 + \beta} \quad (23)$$

$$x_B = \frac{m_{B,t}}{m_t} = \frac{\beta}{1 + \beta}$$

so that

$$C = \ln \frac{1}{1 + \beta} + \beta \ln \frac{\beta}{1 + \beta}$$

and

$$\begin{aligned} \ln x_A + \beta \ln x_B = (1 + \beta) \int_0^m (\alpha_n^{-1} - 1) \frac{dm}{m} + \\ \ln \frac{1}{1 + \beta} + \beta \ln \frac{\beta}{1 + \beta} \end{aligned} \quad (24)$$

Equation 24 is the analog of eq 3 for a two-component system and is applied to molecular weight *vs.* concentration data at constant β . If species B is absent, $\beta = 0$ and eq 24 reduces to eq 3, the parallel equation for a self-associating system. Equation 24 provides the key for the analysis of two-component systems.

Special Cases. It is profitable at this point to consider three special cases, frequently encountered in practice, which the mole fractions of A and B may be obtained directly with the use of eq 24.

(a) If $j \leq 1$, so that no more than one residue of B oc-

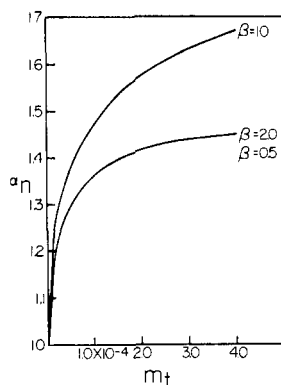
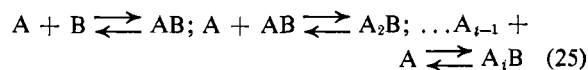


FIGURE 1: The number-average degree of association as a function of m_t for two values of β for the reaction $A + B \rightleftharpoons AB$ ($K = 10^5$; $M_A = M_B = 10^5$).

curs in any complex and A and B do not self-associate, then the only equilibria are



We then have

$$m = [A] + m_{B,t} \quad (26)$$

$$x_A = (m - m_{B,t})/m$$

This may be substituted directly in eq 24, yielding x_B . If x_A and x_B are known, $[A]$ and $[B]$ may be obtained from 20. When $[A]$ and $[B]$ are known, the consecutive association constants may be found from the relations

$$m = [A] + [B] + K_{11}[A][B] + K_{21}[A]^2[B] + \dots \quad (27)$$

$$Y([B]) \equiv \frac{m - [A] - [B]}{[A]}$$

at constant β

$$\left(\frac{dY}{d[B]} \right)_{m=0} = K_{11}$$

$$\left\{ \frac{d}{d[B]} \left(\frac{Y - K_{11}[B]}{[A]} \right) \right\}_{m=0} = K_{21}, \text{ etc.}$$

also

$$\lim_{m \rightarrow 0} \frac{Y}{[B]} = K_{11}, \text{ etc.}$$

(b) If the restriction upon the self-association of A is removed, then in addition to eq 25, we have equilibria of the kind given by eq 1. We then have

$$m = [A] + g([A]) + m_{B,t} \quad (28)$$

where

$$g([A]) = \sum_{i \geq 1} K_{i0}[A]^i$$

when $[B] = 0$

$$= m - [A]$$

By independent measurements upon pure A, $[A]$ may be obtained as a function of $[A] + g([A])$, using eq 3. Then

$$[A] + g([A]) = m - m_{B,t} \quad (29)$$

Using eq 28 and 29, $[A]$ may be obtained. $[B]$ may then be computed from eq 24. Defining

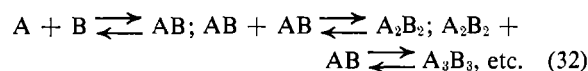
$$Z([B]) = (m - [A] - g[A] - [B])/[A] \quad (30)$$

we have at constant β

$$\left(\frac{dZ}{d[B]} \right)_{m=0} = K_{11} \quad (31)$$

$$\left\{ \frac{d}{d[B]} \left(\frac{Z - K_{11}[B]}{[A]} \right) \right\}_{m=0} = K_{21}, \text{ etc.}$$

(c) If A and B associate to form a dimer, which subsequently associates, so that the only equilibria are



Then, if $\beta = 1$

$$\ln x_A = \ln x_B \quad (33)$$

and from eq 24

$$2 \ln x_A = 2 \int_0^m (\alpha_n^{-1} - 1) \frac{dm}{m} + 2 \ln (1/2) \quad (34)$$

$$\ln (x_A) = \int_0^m (\alpha_n^{-1} - 1) \frac{dm}{m} + \ln (1/2)$$

Since

$$m = [A] + [B] + K_{11}[A][B] + K_{22}[A]^2[B]^2 + \dots \quad (35)$$

$$Y \equiv (m - [A] - [B])/[A]$$

at constant β

$$K_{11} = \left(\frac{dY}{d[B]} \right)_{m=0}$$

$$K_{22} = \left\{ \frac{d}{d[B]} \left[(Y - K_{11}[B])/[A][B] \right] \right\}_{m=0}$$

General Case. We return to the more general case, where there is no restriction upon the number of A or B units in any complex species. The value of [A] or [B] cannot in this case be computed directly from a simple relationship, such as eq 26 or 29. To obtain a second equation, in addition to eq 24, which relates x_A and x_B , measurements may be made for a series of different values of β . In this way a family of curves of α_n vs. m may be obtained for a wide range of β values (Figure 1).

If a vertical line is drawn through the family of curves, corresponding to a particular value of m , a series of points is obtained giving α_n as a function of β at constant m . If the right-hand side of eq 24 is designated as ϕ , there corresponds a value of ϕ for each value of β . The values of ϕ may be computed from eq 24, evaluating the integral by graphical integration.

When m is constant, then a relationship exists of the form

$$m(\ln [A], \ln [B]) = C \quad (36)$$

or

$$m(\ln [A], \ln [B]) - C = 0$$

and

$$\begin{aligned} \frac{d \ln [A]}{d \ln [B]} &= \frac{d \ln x_A}{d \ln x_B} \\ &= - \frac{\partial m / \partial \ln [B]}{\partial m / \partial \ln [A]} \\ &= - \frac{m_{B,t}}{m_{A,t}} = -\beta \end{aligned}$$

Thus, if m is constant and β varies

$$d \ln x_A = -\beta d \ln x_B \quad (m = \text{constant}) \quad (37)$$

$$\begin{aligned} \int d \ln x_A &= - \int \beta d \ln x_B \\ &= -\beta \ln x_B + \int \ln x_B d\beta \end{aligned}$$

If integration is carried out from $\beta = 0$ to β at constant m , we have

$$\ln x_A = -\beta \ln x_B + \int_0^\beta \ln x_B d\beta \quad (38)$$

Introducing eq 38 into 24, where the value of ϕ is evaluated for the given values of β and m

$$\int_0^\beta \ln x_B d\beta = \phi_{m,B} \quad (39)$$

Or, upon differentiating

$$\ln x_B = \frac{d\phi}{d\beta} \quad (40)$$

The value of $\ln x_B$ is thus equal to the slope of the tangent to the curve of ϕ vs. β at the given values of m and β . Combining this with the value of $\ln x_A + \beta \ln x_B$ obtained from eq 24, $\ln x_A$ may be computed for these values of m and β .

If x_A and x_B are both known, [A] and [B] may be obtained. The values of the consecutive association constants may be computed from the equation

$$\begin{aligned} m - [A] - [B] - g([A]) - h([B]) = \\ X([A],[B]) = K_{11}[A][B] + K_{12}[A][B]^2 + \\ K_{21}[A]^2[B] \dots \quad (41) \end{aligned}$$

where

$$g([A]) = \sum_{i \geq 1} K_{i0}[A]^i$$

$$h([B]) = \sum_{j \geq 1} K_{0j}[B]^j$$

The quantities g and h may be obtained as functions of [A] and [B], respectively, by independent measurements upon pure A and B. If A and B do not self-associate, $g = h = 0$. The heterogeneous association constants may be computed from (at constant β)

$$K_{11} = \frac{d}{d[B]} \left(\frac{X}{[A]} \right)_{m=0} \quad (42)$$

$$K_{12} = \frac{d}{d[B]} \left(\frac{X}{[A][B]} \right)_{m=0}, \text{ etc.}$$

To summarize, the computation of a complete set of association constants requires in the most general case: (a) experimental values of α_n as a function of $m (= m_t/\alpha_n)$ for a series of values of β ; (b) computed values of ϕ as a function of β for a series of values of m .

Computation Methods. As a simple example consider the system $A + B \rightleftharpoons AB$, with K_{11} equal to 10^5 . Figures 1 and 2 show α_n as a function of m_t and $(\alpha_n^{-1} - 1)/m$ as a function of m , respectively, for two values of β . Figure 3 shows the variation of ϕ with β for $m = 2.0 \times 10^{-4}$.

The integral

$$\int_0^m \frac{\alpha_n^{-1} - 1}{m} dm$$

for a particular value of β is equal to the areas under the curves of Figure 2. Using curves drawn from six points, the values of ϕ for $m = 2.0 \times 10^{-4}$, as obtained on this basis from eq 24, agree to within 1% with those computed directly.

The estimation of $\ln x_B$ from $d\phi/d\beta$ by purely graphical methods is subject to considerably more error. For

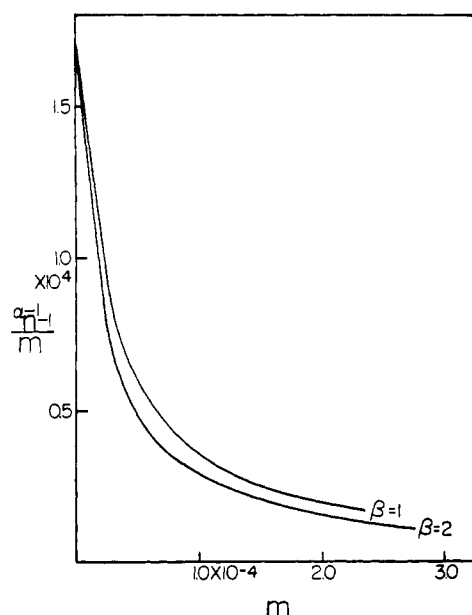


FIGURE 2: Values of $(\alpha_n^{-1} - 1)/m$ as a function of m for the two curves of Figure 1.

the numerical example considered here, the uncertainty in $\ln x_B$, as obtained from the slope of the tangent, is about 5%. For $m = 2 \times 10^{-4}$, $\beta = 1$, the corresponding uncertainty in x_B is about 10%. For actual data, with some scatter, the error may of course be greater.

In practice the method of choice for obtaining $d\phi/d\beta$ would of course be by computer fitting of the experimental values of ϕ vs. β to a polynomial in β . The numerical value of $d\phi/d\beta$ for particular values of β and m can be obtained in this way with considerably more precision, to within 1% in favorable cases.

The final step of evaluating the association constants from the dependence of m upon $[A]$ and $[B]$ may be carried out either by direct graphical methods, or by polynomial fitting, using a computer (Jeffrey and Coates, 1966). If computer facilities are available the latter is probably the method of choice.

Discussion

The theory outlined in the preceding section applies, in principle, to a completely general model, subject to the assumption of ideality. In practice, the latter restriction is rarely a severe drawback, since, at ionic strengths of 0.1 or greater and concentrations less than 10 g/l., deviations from ideality are seldom sufficient to introduce errors of more than a few per cent into the apparent molecular weight, except perhaps under conditions of unusually high net charge per monomer unit. Even in the case of β -lactoglobulin at acid pH (Albright and Williams, 1968), where the magnitude of the second virial coefficient is unusually high, nonideality corrections only become of major importance for concentrations above 5 g/l.

Osmotic pressure offers many advantages for the study of associating protein systems of this kind. In particular, the use of a number-average molecular weight

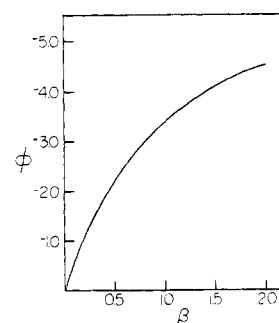


FIGURE 3: The variation of ϕ with β for the system of Figure 1 ($m = 2.0 \times 10^{-4}$).

avoids the undue weighing of the higher complex species characteristic of methods yielding a weight-average molecular weight, as is notoriously the case with light scattering. It should also be stressed that the theory outlined here may be applied without modification to any technique yielding a number-average molecular weight, including the classical colligative methods. It should therefore be applicable to systems of small molecules.

The use of sedimentation equilibrium to study systems of this kind encounters serious difficulties which may preclude a workable analysis except in favorable cases. In particular, if the monomer units differ significantly in molecular weight, a redistribution of monomer species may occur within the cell, so that the ratio of total monomer concentrations is a function of radial position. Since this ratio cannot normally be measured directly, an essential piece of information is lacking.

In practice, of course, the precision falls off sharply for the higher association constants. Until more precise techniques are available it will rarely be feasible to obtain more than the first three or four constants by this kind of approach.

The extension of the basic theory to nonideal systems encounters severe difficulties, which should not however be insuperable. In a subsequent publication it is hoped to make this extension, thereby completing the theory.

References

- Adams, E. T., Jr. (1964), *Proc. Natl. Acad. Sci. U. S.* **51**, 509.
- Adams, E. T., Jr. (1965), *Biochemistry* **4**, 1646.
- Adams, E. T., Jr., and Fujita, H. (1963), in *Ultracentrifugal Analysis in Theory and Experiment*, Williams, J. W., Ed., New York, N. Y., Academic.
- Adams, E. T., Jr., and Williams, J. W. (1964), *J. Am. Chem. Soc.* **86**, 3454.
- Albright, D. A., and Williams, J. W. (1968), *Biochemistry* **7**, 67.
- Jeffrey, P. D., and Coates, J. H. (1966), *Biochemistry* **5**, 489.
- Reithel, F. J. (1963), *Advan. Protein Chem.* **18**, 123.
- Steiner, R. F. (1952), *Arch. Biochem. Biophys.* **39**, 333.
- Steiner, R. F. (1954), *Arch. Biochem. Biophys.* **49**, 400.
- Ts'o, P. O. P., Melvin, I. S., and Olson, A. C. (1963), *J. Am. Chem. Soc.* **85**, 1289.