

# VIRIAL EXPANSIONS FOR IDEAL SELF-ASSOCIATING SYSTEMS

W. F. STAFFORD III *and* D. A. YPHANTIS

*From the Biochemistry and Biophysics Section, Biological Sciences Group and Institute of Materials Science, The University of Connecticut, Storrs, Connecticut 06268*

**ABSTRACT** Theoretical virial equations for self-associating systems governed by mass action have been derived assuming the solute to be ideal except for this solute-solute interaction. In particular, monomer-polymer association involving two molecular species and isodesmic association involving an indefinite number of molecular species have been treated analytically. The usefulness of such virial equations is severely limited by their extremely narrow interval of convergence.

## INTRODUCTION

It is common practice when considering colligative properties of nonelectrolyte solutions to express the osmotic pressure (1-3) and the excess chemical potential of the solvent (4) or solute (5) by means of a power series expansion in the solute concentration. Such virial equations have proven to be of much value for the analysis and interpretation of experimental data.

In particular, it has been assumed (6) that nonideality attributable to self-association could be treated with a virial equation. We show here, however, that such treatment is seriously limited by the extremely narrow interval of convergence for such series. We consider two types of self-association: (a) a monomer-polymer equilibrium in which only two macromolecular species are present and (b) monomer-polymer equilibria of indefinite extent in which the free energy change for the sequential addition of each monomer is constant. More complex cases appear difficult to treat analytically but are expected to yield similarly small radii of convergence.

The usual virial equation for the osmotic pressure  $\Pi$  defines the colligative virial coefficients  $B_n$  through the relationship

$$\Pi/cRT = (1/M + B_2c + B_3c^2 + B_4c^3 + \dots), \quad (1)$$

where  $R$  is the gas constant,  $T$  the absolute temperature,  $M$  the molecular weight of the solute, and  $c$  the concentration of solute in grams  $\times$  centimeters<sup>-3</sup>. The activity coefficient  $\gamma$  of the solute may be expressed with the same colligative virial coefficients

(7) as

$$\ln y = 2MB_2c + (3/2)MB_3c^2 + (4/3)MB_4c^3 + \dots \quad (2)$$

The apparent weight-average molecular weight  $M_w(c)$  for a nonideal solute component may be written as

$$M_w(c) = \frac{M}{1 + c \left( \frac{\partial \ln y}{\partial c} \right)_{P, T}} = \frac{1}{1/M + 2B_2c + 3B_3c^2 + 4B_4c^3 + \dots}, \quad (3)$$

where the right-hand equality follows on substitution of equation 2. We shall use this expression, equation 3, as our starting point, assuming that the solute is a single thermodynamic component and that all the nonideality arises from self-association of this component as specified by the mass action law. Thus, we specifically exclude any other source of nonideality, such as Donnan effects, excluded volume effects and electrostatic interactions, etc., so that the virial equations obtained represent only the self-association. The inclusion of further nonideality, such as from the Donnan effect in self-associating systems (8), should not materially affect the present conclusions.

#### MONOMER $N$ -MER AS A FUNCTION OF TOTAL CONCENTRATION

The power series expansion of  $(1/M_w)$  for a monomer  $N$ -mer self-associating system was obtained by calculating the necessary derivatives from an implicit relationship between  $(1/M_w)$  and total macromolecular solute concentration. This relationship was obtained in the following way.

The weight-average molecular weight is defined in terms of the weight fraction of  $N$ -mer in the usual way:

$$M_w = M_1(1 - \alpha) + NM_1\alpha, \quad (4)$$

where  $\alpha$  is the weight fraction of polymer,  $N$  is the degree of polymerization and  $M_1$  is the monomer molecular weight. Then a function  $f(\alpha)$  is defined as:

$$f(\alpha) \equiv \frac{M_1}{M_w} = \frac{1}{1 + (N - 1)\alpha}. \quad (5)$$

We then write the mass action law for a monomer  $N$ -mer self-association as:

$$K = \frac{c_N}{c_1^N} = \frac{\alpha}{(1 - \alpha)^N c^{N-1}}, \quad (6)$$

where  $c_1 = (1 - \alpha)c$ , and  $c_N = \alpha c$ , with the total macromolecular solute concentration in grams  $\times$  centimeters<sup>-3</sup> denoted by  $c$ .

We combine equations 5 and 6 to eliminate  $\alpha$  and obtain:

$$q \equiv Kc^{(N-1)} = \frac{(1/f - 1)(N - 1)^{(N-1)}}{(N - 1/f)^N}. \quad (7)$$

Letting

$$\phi = K[c/(N - 1)]^{(N-1)}, \quad (8)$$

we obtain the implicit relationship between  $f$  and  $\phi$ :

$$\phi = \frac{(1/f - 1)}{(N - 1/f)^N}. \quad (9)$$

We now have an implicit functional relationship between  $(1/M_w)$  and  $c$  in terms of  $f$  and  $\phi$ . The logical variable in which to expand  $f$  appears to be  $\phi$  since only integral powers of  $c^{N-1}$  have nonzero derivatives (see the Appendix).

### COMPUTATION OF SUCCESSIVE DERIVATIVES

The necessary derivatives were obtained with a computer program written in Formac, a PL/1-based computer language that enables the manipulation of symbolic algebraic expressions. An algorithm was devised to calculate successive total derivatives of one variable with respect to another having an implicit functional relationship. In this case, for example, the implicit relationship was between  $f$  and  $\phi$ .

The first derivative was obtained by implicit differentiation of expression 9. If  $G^i(\phi, f)$  represents the  $i$ th order total derivative of  $f$  with respect to  $\phi$ , the  $(i + 1)$ th order total derivative was obtained from the following general recursion relation:

$$G^{i+1}(\phi, f) = \frac{dG^i(\phi, f)}{d\phi} = \frac{\partial G^i(\phi, f)}{\partial \phi} + \frac{\partial G^i(\phi, f)}{\partial f} \frac{df}{d\phi}, \quad (10)$$

where

$$G^1(\phi, f) = df/d\phi.$$

The following expression for the  $i$ th order derivative evaluated at  $\phi = 0, f = 1$  was inferred from an analysis of the computer-generated derivatives:

$$\begin{aligned} \left(\frac{d^i f}{d\phi^i}\right)_{\phi=0} &= (-1)^j \frac{(jN - 1)!}{[jN - (1 + j)]!} (N - 1)^{(N-1)j}, \\ &= (-1)^j \left[ \prod_{i=1}^j (jN - i) \right] (N - 1)^{(N-1)j}. \end{aligned} \quad (11)$$

The necessary derivatives for the expansion of  $M_1/M_w$  vs.  $q = Kc^{N-1}$  are given by:

$$\begin{aligned} \left(\frac{d^i f}{dq^i}\right)_{q=0} &= (-1)^j \frac{(jN - 1)!}{[jN - (1 + j)]!}, \\ &= (-1)^j (jN - 1)(jN - 2)(jN - 3) \cdots (jN - j), \end{aligned} \quad (12)$$

since  $d\phi/dq = (N - 1)^{(1-N)}$ , and all  $d^i \phi/dq^i = 0, i \geq 2$ .

Thus, the virial equation for the weight-average molecular weight of an ideal self-associating system assumes the following form:

$$\frac{1}{M_w} = \frac{1}{M_1} \left\{ 1 + \sum_{j=1}^{\infty} (-1)^j \frac{(jN - 1)!}{j! [jN - (1 + j)]!} q^j \right\}, \quad (13 a)$$

$$\frac{1}{M_w} = \frac{1}{M_1}$$

$$\cdot \left[ 1 - (N - 1)Kc^{N-1} + \frac{1}{2} (2N - 2)(2N - 1)K^2 c^{2N-2} - \dots \right]. \quad (13 b)$$

For example, the expression for a monomer-dimer equilibrium is

$$\frac{1}{M_w} = \frac{1}{M_1} (1 - Kc + 3K^2 c^2 - \dots),$$

for monomer-trimer,

$$\frac{1}{M_w} = \frac{1}{M_1} (1 - 2Kc^2 + 10K^2 c^4 - \dots),$$

TABLE I  
VALUES OF THE WEIGHT FRACTION OF POLYMER  $\alpha$  AND THE  
NORMALIZED WEIGHT-AVERAGE MOLECULAR WEIGHT ( $M_w/M_1$ )  
AT THE RADIUS OF CONVERGENCE  $Kc_R^{N-1}$  FOR VARIOUS  
DEGREES  $N$  OF POLYMERIZATION

$N$	$Kc_R^{N-1}$	$\alpha_R$	$(M_w/M_1)_R$
2	0.250	0.1716	1.172
3	0.148	0.1069	1.212
4	0.106	0.0767	1.230
5	0.0819	0.0601	1.240
6	0.0669	0.0494	1.247
7	0.0567	0.0420	1.252
8	0.0491	0.0365	1.256
9	0.0434	0.0323	1.258
10	0.0387	0.0289	1.260
20	0.0189	0.0142	1.270
40	0.0093	0.0070	1.273
80	0.0046	0.0035	1.277
160	0.0024	0.0018	1.278

limit  $(M_w/M_1)_R = 1.2785$   
( $N \rightarrow \infty$ )

and for monomer-tetramer,

$$\frac{1}{M_w} = \frac{1}{M_1} (1 - 3Kc^3 + 21K^2c^6 - \dots).$$

The radius of convergence (9) of the series in terms of  $q$  is given by

$$q_R = \lim_{j \rightarrow \infty} |a_j/a_{j+1}|,$$

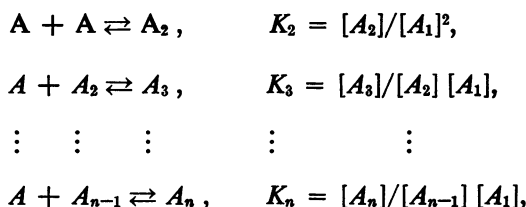
where the  $a_j$  are the coefficients of the series. The value of  $q_R$  for the series of equation 13 *a* can be shown to be  $q_R = (N - 1)^{N-1}/N^N$ . Thus, the corresponding expansion in terms of  $c$ , equation 13 *b*, must converge for all  $c < c_R$  where  $c_R$  is given by

$$c_R = (q_R/K)^{1/N-1} = (N - 1) (KN^N)^{1/1-N}. \quad (14)$$

Table I presents values of  $\alpha$ , the weight fraction of polymer, and of  $M_w/M_1$  for various degrees of polymerization  $n$  at the radius of convergence. From this it can be seen that the region of convergence is, in general, quite limited;

#### INDEFINITE ASSOCIATING SYSTEM

In an indefinite associating system, for which successive molar equilibrium constants are equal (the isodesmic model), we have



where  $K_i = K$ , and  $[A_i]$  is the concentration of species  $A_i$  in moles per liter;

As above, we define a function  $f(c)$  as:

$$f(c) \equiv \frac{M_1}{M_w}, \quad (15)$$

and obtain from van Holde et al. (10) the relation

$$f(c) = (1 + kc)^{-1/2}. \quad (16)$$

Where  $k = 4 \times 10^3 K/M_1$ ,  $M_1$  is the monomer molecular weight and  $c$  is the total macromolecular solute concentration in grams  $\times$  centimeters<sup>-3</sup>.

The successive derivatives are easily obtained without the aid of the high speed

electronic computer by ordinary straightforward means:

$$\left(\frac{d^j f}{dc^j}\right)_{c=0} = \left(-\frac{1}{2}\right)^j \prod_{i=1}^j (2i - 1). \quad (17)$$

The power series expansion is written as:

$$1/M_w = 1/M_1 \left[ 1 + \sum_{j=1}^{\infty} \left[ \left(-\frac{1}{2}\right)^j (kc)^j \right] (1/j!) \prod_{i=1}^j (2i - 1) \right]. \quad (18)$$

This series has a radius of convergence  $Kc/M_1 \leq 2.5 \times 10^{-4}$  such that for convergence  $c < 2.5 \times 10^{-4}(M_1/K)$ . At the radius of convergence the normalized weight-average molecular weight ( $M_w/M_1$ ) is  $\sqrt{2}$  and the *weight* fraction of monomer is 0.68629.

## DISCUSSION

The uniqueness of the above virial equations is assured by a theorem for power series which states that "a function  $f(x)$  can be represented by a power series in  $x$  in only one way, if at all (11)." The surprisingly narrow interval of convergence of these series clearly dictates the exercise of caution in attempts to use untruncated virial expressions in the analysis of self-associating systems. Caution should also be exercised in attempts to obtain meaningful estimates of equilibrium constants by taking the limiting slope at infinite dilution, since the effective linear range of  $M_1/M_w$  as a function of  $c^{N-1}$  at best extends to 7% association, a range in which experimental error usually is a large fraction of the change in molecular weight.

## APPENDIX

We wish to show that only integral powers of  $c^{N-1}$  need be considered, and hence that equation 19 is true:

$$\left(\frac{d^j f}{dc^j}\right)_{c=0} \neq 0 \quad \text{only for} \quad j = (N - 1)i, \quad i = 1, 2, 3, 4, \dots \quad (19)$$

This can be shown in the following way: direct differentiation of equation 8 gives

$$\left(\frac{d^i \phi}{dc^i}\right)_{c=0} = (N - i)!(N - 1)^{(N-1)} Kc^{(N-i-1)}.$$

Thus, the derivatives of equation 19 are nonzero only when  $(N - 1) = i$ .

Therefore,

$$\frac{df}{dc} = \frac{df}{d\phi} \frac{d\phi}{dc} \neq 0 \quad \text{only for} \quad N = 2,$$

$$\frac{d^2f}{dc^2} = \frac{d^2f}{d\phi^2} \left(\frac{d\phi}{dc}\right)^2 + \frac{df}{d\phi} \frac{d^2\phi}{dc^2} \neq 0 \quad \text{only for } N = 2 \text{ and } N = 3,$$

$$\frac{d^3f}{dc^3} = \frac{d^3f}{d\phi^3} \left(\frac{d\phi}{dc}\right)^3 + 3 \frac{d^2f}{d\phi^2} \frac{d^2\phi}{dc^2} \frac{d\phi}{dc} + \frac{df}{d\phi} \frac{d^3\phi}{dc^3} \neq 0 \quad \text{only for } N = 2 \text{ and } N = 4,$$

$$\frac{d^4f}{dc^4} \neq 0 \quad \text{only for } N = 2, \quad N = 3, \text{ and } N = 5.$$

and so on.

This work was supported by grants No. GB-13790 and GB-30825x from the National Science Foundation and by U. S. Public Health Service training grant No. GM-00317.

Computer time was made available by the University of Connecticut Computing Center which is supported, in part, by grant No. GJ-9 from the National Science Foundation.

*Received for publication 21 February 1972.*

## REFERENCES

1. McMILLAN, W. G., and J. E. MAYER. 1945. *J. Chem. Phys.* **14**:276.
2. ISHIHARA, A. 1951. *J. Chem. Phys.* **19**:397.
3. STIGTER, D. 1964. *J. Phys. Chem.* **64**:118.
4. KOZAK, J. J., W. S. KNIGHT, and W. KAUZMANN. 1968. *J. Chem. Phys.* **48**:675.
5. HILL, T. L. 1960. *Introduction to Statistical Thermodynamics*. Addison-Wesley Publishing Co., Inc., Reading, Mass.
6. ROARK, D. E., and D. A. YPHANTIS. 1969. *Ann. N. Y. Acad. Sci.* **164**:255.
7. BALDWIN, R. L., and K. E. VAN HOLDE. 1960. *Fortschr. Hochpolym.-Forsch.* **1**:451.
8. ROARK, D. E., and D. A. YPHANTIS. 1971. *Biochemistry*. **10**:3241.
9. KAPLAN, W. 1959. *Advanced Calculus*. Addison-Wesley Publishing Co., Inc., Reading, Mass.
10. VAN HOLDE, K. E., G. P. ROSSETTI, and R. D. DYSON. 1969. *Ann. N. Y. Acad. Sci.* **164**:279.
11. COURANT, R. 1959. *Differential and Integral Calculus*. Interscience Publishers Inc., New York. **1**.