Excess Free Energy Approach to the Estimation of Solubility in Mixed Solvent Systems I: Theory

N. A. WILLIAMS * and G. L. AMIDON *ˣ

Abstract □ An approach is developed by which the solubility of an organic compound in mixed solvents may be estimated. In this approach, an expression for the excess Gibbs free energy of mixing for multicomponent solvent systems was used to obtain parameters characteristic of the interaction between the solvents. A fairly simple equation which predicts the solubility of a solute in a binary solvent system over the entire solvent composition range was then derived. The equation may be partitioned into terms that contain (a) pure solvent solubilities, (b) solvent–solvent interaction contributions, and (c) contributions from the solute–mixed solvent interactions. The required data are the molar volume of the solute, the pure solvent solubilities, and, theoretically, one experimentally determined solubility in a solvent mixture. The equation can be easily extended for systems with three or more solvents.

Keyphrases □ Solubility—mixed solvent systems, estimation by an excess free energy approach, theory □ Excess free energy—use in estimating solubility in mixed solvent systems, theory □ Mixed solvent systems—estimation of solubility, excess free energy approach, theory

For pharmaceutical purposes, it is often necessary to dissolve a nonpolar or slightly polar drug in a mixture of water and one or more cosolvents such as ethanol, glycerol, propylene glycol, and low-molecular weight polyethylene glycols to increase its solubility. It is also often desirable to know if and where a maximum exists in the solubility profile of the drug in the mixture of solvents. Apart from determining such a profile experimentally over the whole solvent composition range, no general method exists at present which can, with a minimum of experiments, completely describe the solubility in the solvent mixture. Typical approaches to the estimation of solubility in mixed solvents merely express solubility in terms of a power series in mole or volume fraction multiplied by arbitrary numbers that have little physical meaning (1). Although useful such equations are by their nature restricted to and only useful for the particular solute–solvent systems for which they were formulated.

Recently Martin and coworkers (2–5) have attempted to use an extended form of the regular solution equation for the solubility of solids to estimate solubility in pure and mixed solvents. This is accomplished, in effect, by adding a term, W, to the expression for the regular solution activity coefficient to account for non-dispersion-type solute–solvent interactions. W is then estimated from differences between experimental and calculated solubilities using the regular solution equation. While their method is an improvement over the purely empirical approach, it does have a few disadvantages. For instance, it requires obtaining the heat of fusion, melting point, molar volume, and solubility parameter of the solute and the solubility parameter of the solvent or solvent mixture. When these data are not readily available, the experimental labor involved in determining them may be such as to frustrate the ultimate aim of the approach, *i.e.*, to estimate solubility with a minimal number of experiments. As the authors have also pointed out (2), the correction term (W) is difficult to determine independently, and the solubility parameter of the solute may assume a different value in a different solvent system (although this does not seriously affect the results) (4).

This paper presents the theoretical aspects of a method based on an excess free energy model which can be used to characterize the solubility of compounds in binary and ternary solvent systems. The method is general and applicable, in principle, to virtually any solute–mixed solvent system as long as the solubility is not high. The data required are the molar volume of the solute, the pure solvent solubilities, and, theoretically, one experimentally determined solubility in a solvent mixture. The last item is required for estimating C_2, a ternary interaction term. Although estimated empirically, C_2, unlike W, results logically from the development of the excess free energy model for the system. Subsequent papers in this series (6, 7) discuss the usefulness of this approach in describing experimental data from the literature.

THEORETICAL

For a solute in solution in equilibrium with its solid phase:

$$f_2^{soln} = f_2^{pure} \qquad \text{(Eq. 1)}$$

where f_2^{soln} is the fugacity of the solute in solution and f_2^{pure} is the fugacity of the pure solid. The fugacity of the solute in solution may be expressed as (8):

$$f_2^{soln} = x_2 \gamma_2 f_2^0 \qquad \text{(Eq. 2)}$$

where x_2 is the mole fraction concentration of the solute, γ_2 is the symmetric convention activity coefficient, and f_2^0 is the fugacity of the hypothetical pure liquid at the same temperature and pressure as the solution (as is shown later, f_2^0 drops out of the equation and, therefore, its evaluation does not present a problem).

The symmetric convention activity coefficient (γ_2) has the property that $\gamma_2 \to 1$ as $x_2 \to 1$, while the unsymmetric convention activity coefficient (γ_2^*) has the property that $\gamma_2^* \to 1$ as $x_2 \to 0$. The relationship between the two activity coefficients has been derived by Prigogine and Defay (9):

$$\frac{\gamma_2}{\gamma_2^*} = \lim_{x_2 \to 0} \gamma_2 \qquad \text{(Eq. 3)}$$

If x_2 approaches zero, γ_2^* is equal to ~1 and Eq. 3 becomes:

$$\gamma_2 \approx \lim_{x_2 \to 0} \gamma_2 \qquad \text{(Eq. 4)}$$

This situation is described by Henry's Law which may be written as:

$$\lim_{x_2 \to 0} \frac{f_2^{soln}}{x_2} = H_2 \qquad \text{(Eq. 5)}$$

where H_2 is the Henry's Law constant.

If we assume that the mole fraction solubility is sufficiently small such that Henry's Law holds up to the solubility limit, we can write:

$$f_2^{soln} = f_2^{pure} = H_2 x_2^s \qquad \text{(Eq. 6)}$$

where x_2^s is the mole fraction solubility. In many cases of interest, the mole fraction solubility of the drug or chemical in question is low enough even in the pure cosolvent to justify such an assumption.

Comparison of Eqs. 2 and 5 shows that:

$$\lim_{x_2 \to 0} \gamma_2 = \frac{H_2}{f_2^0} \qquad \text{(Eq. 7)}$$

If we take the logarithm of Eq. 7 and then write the expression for the solute in solvents 1 and 3 separately, we obtain for solvent 1:

$$\lim_{x_2 \to 0} \ln \gamma_{2,1} = \ln \frac{H_{2,1}}{f_2^0} \qquad \text{(Eq. 8a)}$$

and for solvent 3:

$$\lim_{x_2 \to 0} \ln \gamma_{2,3} = \ln \frac{H_{2,3}}{f_2^0} \qquad \text{(Eq. 8b)}$$

We note here that subscript 2 is used for the solute while subscripts 1, 3, 4, *etc.* are used for the pure solvents. For a binary mixture (denoted by subscript m) of solvents 1 and 3, Eq. 7 becomes:

$$\lim_{x_2 \to 0} \ln \gamma_{2,m} = \ln \frac{H_{2,m}}{f_2^0} \qquad \text{(Eq. 8c)}$$

If we take the logarithms of Eqs. 1, 2, and 4 and combine them, we obtain, at the solubility limit:

$$\ln x_2^s = \ln \left(\frac{f_2^{\text{pure}}}{f_2^0} \right) - \lim_{x_2 \to 0} \ln \gamma_2 \qquad \text{(Eq. 9)}$$

Substituting the expressions for $\ln \gamma_2$ in Eqs. 8a–c we obtain, for solvent 1:

$$\ln x_{2,1}^s = \ln \frac{f_2^{\text{pure}}}{H_{2,1}} \qquad \text{(Eq. 10a)}$$

for solvent 2:

$$\ln x_{2,3}^s = \ln \frac{f_2^{\text{pure}}}{H_{2,3}} \qquad \text{(Eq. 10b)}$$

and for a mixture of solvents 1 and 3:

$$\ln x_{2,m}^s = \ln \frac{f_2^{\text{pure}}}{H_{2,m}} \qquad \text{(Eq. 10c)}$$

We now define the volume fraction z_i of any component i in the mixture as:

$$z_i = \frac{x_i q_i}{\sum_i x_i q_i} \qquad \text{(Eq. 11)}$$

where q_i is the molar volume of component i. If we let x_2 approach zero, z_2 also approaches zero, and z_1 and z_3 become solute-free volume fractions, *i.e.*:

$$\hat{z}_1 + \hat{z}_3 = 1 \qquad \text{(Eq. 12)}$$

where the circumflex denotes solute-free terms. From Eq. 6 we can write:

$$\ln f_2^{\text{pure}} = \ln x_{2,1}^s + \ln H_{2,1} \qquad \text{(Eq. 13a)}$$

and

$$\ln f_2^{\text{pure}} = \ln x_{2,3}^s + \ln H_{2,3} \qquad \text{(Eq. 13b)}$$

If we multiply Eq. 13a by $\hat{z}_1$, Eq. 13b by $\hat{z}_3$, and add them, we obtain:

$$\ln f_2^{\text{pure}} = \hat{z}_1 (\ln x_{2,1}^s + \ln H_{2,1}) + \hat{z}_3 (\ln x_{2,3}^s + \ln H_{2,3}) \qquad \text{(Eq. 14)}$$

Eq. 14 substituted into Eq. 10c gives:

$$\ln x_{2,m}^s = \hat{z}_1 \ln x_{2,1}^s + \hat{z}_3 \ln x_{2,3}^s - (\ln H_{2,m} - \hat{z}_1 \ln H_{2,1} - \hat{z}_3 \ln H_{2,3}) \qquad \text{(Eq. 15)}$$

Equation 15 shows that the solubility of a compound in a binary solvent mixture is a volume fraction-weighted sum of the two pure solvent solubilities plus what can be termed an excess Henry's Law constant H_2^{Ex}. In general, we may write:

$$\ln x_{2,m}^s = \sum_{\text{solv } i} \hat{z}_1 \ln x_{2,i}^s - \ln H_2^{\text{Ex}} \qquad \text{(Eq. 16)}$$

$$\ln H_2^{\text{Ex}} = \ln H_{2,m} - \sum_{\text{solv } i} \hat{z}_1 \ln H_{2,i} \qquad \text{(Eq. 17)}$$

For Eqs. 15–17 to be useful, an expression for $\ln H_2^{\text{Ex}}$ must be developed. To obtain such an expression, we have employed an excess free energy model as proposed by Wohl (10).

Excess Free Energy Model—The total free energy (G) of a homo-geneous system consisting of r components may be written as:

$$G = \sum_{i=1}^r n_i g_i^0 + RT \sum_{i=1}^r n_i \ln x_i + \left(\sum_{i=1}^r n_i \right) g^{\text{E}} \qquad \text{(Eq. 18)}$$

where n_i represents the number of moles of the ith component, x_i is the mole fraction of the ith component, and g^{E} is the excess free energy per mole. The first term represents the free energies of each of the pure components (g_i^0 being the free energy per mole of pure component i). The second term is the free energy of ideal mixing, and the third term represents the excess free energy arising from nonidealities in the mixing. The partial molar free energy of component j at constant temperature, pressure, and composition is given by:

$$\left(\frac{\partial G}{\partial n_j} \right)_{T,P,n_{i \neq j}} = g_j = g_j^0 + RT \ln x_j + RT \ln \gamma_j \qquad \text{(Eq. 19)}$$

where

$$\ln \gamma_j = \left[\frac{\partial \left[\sum_{i=1}^r n_i \dfrac{g^{\text{E}}}{RT} \right]}{\partial n_j} \right]_{T,P,n_{i \neq j}} \qquad \text{(Eq. 20)}$$

In 1946, Wohl (10) proposed a general method for expressing the excess free energies of mixtures in terms of increasing powers of the volume fractions (z) of the components. Thus, the excess free energy of a ternary system may be expressed as:

$$\frac{g^{\text{E}}}{RT(x_1 q_1 + x_2 q_2 + x_3 q_3)} = 2a_{12} z_1 z_2 + 2a_{13} z_1 z_3 + 2a_{23} z_2 z_3$$
$$+ 3a_{112} z_1^2 z_2 + 3a_{122} z_1 z_2^2 + 3a_{113} z_1^2 z_3 + 3a_{133} z_1 z_3^2$$
$$+ 3a_{223} z_2^2 z_3 + 3a_{233} z_2 z_3^2 + 6a_{123} z_1 z_2 z_3 + 4a_{1112} z_1^3 z_2$$
$$+ 6a_{1122} z_1^2 z_2^2 + 4a_{1222} z_1 z_2^3 + 4a_{1113} z_1^3 z_3 + 6a_{1133} z_1^2 z_3^2$$
$$+ 4a_{1333} z_1 z_3^3 + 4a_{2223} z_2^3 z_3 + 6a_{2233} z_2^2 z_3^2 + 4a_{2333} z_2 z_3^3$$
$$+ 12a_{1123} z_1^2 z_2 z_3 + 12a_{1223} z_1 z_2^2 z_3 + 12a_{1233} z_1 z_3^2 \qquad \text{(Eq. 21)}$$

where $z_1 = x_1 q_1 / (x_1 q_1 + x_2 q_2 + x_3 q_3 \dots)$ (see Eq. 11). The q values are usually considered a measure of the sizes of the molecules. In this report we have used molar volumes for the q-values; Wohl used the term "four-suffix q equation" in describing Eq. 21 to distinguish it from Margules-type equations which assume that all the molecules are of equal size. The "four" refers to the largest number of subscripts on an a term in the equation. As an example, a_{12} is a constant which represents the interaction between one molecule of component 1 and another of component 2. Also, as Prausnitz (8) has pointed out, the probability that any pair of two molecules consists of one molecule each of components 1 and 2 is given by $2z_1 z_2$. Similarly a_{1133} is the interaction parameter for four interacting molecules two of which are components 1 and 3, respectively, and $6z_1^2 z_3^2$ is the probability that any four neighboring molecules consist of two molecules each of components 1 and 3. Although the a values may be regarded, in a rough sense, as similar to virial coefficients, they do not have a rigorous theoretical basis (8). Their importance lies in the fact that they are physically meaningful parameters.

Higher- or lower-suffix q equations may also be written for the excess free energy depending on the degree of complexity required. A three-suffix equation contains only the first 10 terms on the right-hand side of Eq. 21, and a two-suffix equation contains only the first three terms.

Wohl (10) defined:

$$A_{2-1} = q_2 (2a_{12} + 3a_{112} + 4a_{1112}) \qquad \text{(Eq. 22a)}$$

$$A_{1-2} = q_1 (2a_{12} + 3a_{122} + 4a_{1222}) \qquad \text{(Eq. 22b)}$$

$$A_{3-1} = q_3 (2a_{13} + 3a_{113} + 4a_{1113}) \qquad \text{(Eq. 22c)}$$

$$A_{1-3} = q_1 (2a_{13} + 3a_{133} + 4a_{1333}) \qquad \text{(Eq. 22d)}$$

$$A_{3-2} = q_3 (2a_{23} + 3a_{223} + 4a_{2223}) \qquad \text{(Eq. 22e)}$$

$$A_{2-3} = q_2 (2a_{23} + 3a_{233} + 4a_{2333}) \qquad \text{(Eq. 22f)}$$

$$C_1 = q_1 [(3a_{112} + 3a_{133} + 3a_{223} - 6a_{123}) + 4a_{1112} + 6a_{1122}$$
$$- 4a_{1222} + 6a_{1133} + 4a_{2223} - 12a_{1123}] \qquad \text{(Eq. 22g)}$$

$$C_2 = q_2 [(3a_{112} + 3a_{133} + 3a_{223} - 6a_{123}) + 4a_{2223} + 6a_{2233}$$
$$- 4a_{2333} + 6a_{1122} + 4a_{1333} - 12a_{1223}] \qquad \text{(Eq. 22h)}$$

$$C_3 = q_3 [(3a_{112} + 3a_{133} + 3a_{223} - 6a_{123}) + 4a_{1333} + 6a_{1133}$$

$$- 4a_{1113} + 6a_{2233} + 4a_{1112} - 12a_{1233}] \qquad \text{(Eq. 22i)}$$

$$D_{12} = q_1[4a_{1112} + 4a_{1222} - 6a_{1122}] \qquad \text{(Eq. 22j)}$$

$$D_{13} = q_3[4a_{1333} + 4a_{1113} - 6a_{1133}] \qquad \text{(Eq. 22k)}$$

$$D_{23} = q_2[4a_{2223} + 4a_{2333} - 6a_{2233}] \qquad \text{(Eq. 22l)}$$

If we introduce Eqs. 22a–l into Eq. 21 and differentiate the total excess free energy with respect to n_2, Eq. 23 is obtained:

$$\left[\frac{\partial n_T \frac{g^E}{RT}}{\partial n_2} \right]_{T,P,n_1,n_3} = \ln \gamma_{2,m}$$

$$= z_3^2 \left[A_{2\text{-}3} + 2z_2 \left(A_{3\text{-}2} \frac{q_2}{q_3} - A_{2\text{-}3} - D_{23} \right) + 3z_2^2 D_{23} \right]$$

$$+ z_1^2 \left[A_{2\text{-}1} + 2z_2 \left(A_{1\text{-}2} \frac{q_2}{q_1} - A_{2\text{-}1} - D_{12} \frac{q_2}{q_1} \right) + 3z_2^2 D_{12} \frac{q_2}{q_1} \right]$$

$$+ z_1 z_3 \left[A_{3\text{-}2} \frac{q_2}{q_3} + A_{2\text{-}1} - A_{1\text{-}3} \frac{q_2}{q_1} + 2z_2 \left(A_{1\text{-}2} \frac{q_2}{q_1} - A_{2\text{-}1} \right) \right]$$

$$+ 2z_1 \left(A_{1\text{-}3} \frac{q_2}{q_1} - A_{3\text{-}1} \frac{q_2}{q_3} \right) - 3z_1 z_3 D_{13} \frac{q_2}{q_3} - C_2 z_2 (2 - 3z_2)$$

$$- C_3 \frac{q_2}{q_3} z_3 (1 - 3z_2) - C_1 \frac{q_2}{q_1} z_1 (1 - 3z_2) \right] \qquad \text{(Eq. 23)}$$

where n_T is the total number of moles. Since, in our convention, subscripts 1, 2, and 3 represent one solvent, the solute, and the other solvent, respectively, in the mixture m, $\gamma_{2,m}$ is the activity coefficient of a solute in a mixture of two solvents in which the solute is dissolved. If desired, the activity coefficients of each of the solvents may be obtained by appropriately interchanging the subscripts (10), but they are not needed in this analysis.

For the purpose of our analysis Eq. 23 may be simplified by making the following approximations:

$$3a_{122} \approx 3a_{112}, \quad 4a_{1222} \approx 4a_{1112} \qquad \text{(Eq. 24a)}$$

$$3a_{233} \approx 3a_{223}, \quad 4a_{2333} \approx 4a_{2223} \qquad \text{(Eq. 24b)}$$

With these approximations:

$$\frac{A_{1\text{-}2}}{A_{2\text{-}1}} = \frac{q_1}{q_2}, \quad \frac{A_{2\text{-}3}}{A_{3\text{-}2}} = \frac{q_2}{q_3} \qquad \text{(Eq. 25)}$$

Equations 24a and b imply that we are neglecting three (or more)-body interactions between the solute and each of the solvents. This approximation seems reasonable in view of the fact that we are dealing in most cases with low concentrations of solute. If we apply the same approximation to solvent–solvent interactions, Eq. 23 reduces to the van Laar equations which have been used to fit activity coefficient data with some success.

Substituting Eq. 25 into Eq. 23 and letting x_2 (and hence z_2) approach zero, we obtain:

$$\lim_{x_2 \to 0} \ln \gamma_{2,m} = A_{2\text{-}3} \hat{z}_3 + A_{2\text{-}1} \hat{z}_1 + A_{1\text{-}3} \hat{z}_1 \hat{z}_3 (2\hat{z}_1 - 1) \frac{q_2}{q_1}$$

$$- A_{3\text{-}1} 2\hat{z}_1^2 \hat{z}_3 \frac{q_2}{q_3} - D_{13} 3\hat{z}_1^2 \hat{z}_3^2 \frac{q_2}{q_3} - C_3 \hat{z}_1 \hat{z}_3^2 \frac{q_2}{q_3} - C_1 \hat{z}_1^2 \hat{z}_3 \frac{q_2}{q_1} \qquad \text{(Eq. 26)}$$

where $A_{2\text{-}3}$ is the constant arising from the interaction of the solute with pure solvent 3, and $A_{2\text{-}1}$ is the corresponding constant for the solute and pure solvent 1.

Comparison of Eqs. 8a and b with the first two terms in Eq. 26 shows that:

$$\lim_{x_2 \to 0} \ln \gamma_{2,1} = \ln \frac{H_{2,1}}{f_2^0} = A_{2\text{-}1} \qquad \text{(Eq. 27a)}$$

and

$$\lim_{x_2 \to 0} \ln \gamma_{2,3} = \ln \frac{H_{2,3}}{f_2^0} = A_{2\text{-}3} \qquad \text{(Eq. 27b)}$$

Substituting Eqs. 8c, 27a, and 27b into Eq. 26, we obtain:

$$\ln H_{2,m} = \hat{z}_1 \ln H_{2,1} + \hat{z}_3 \ln H_{2,3} + A_{1\text{-}3} \frac{q_2}{q_1} \hat{z}_1 \hat{z}_3 (2\hat{z}_1 - 1)$$

$$- A_{3\text{-}1} \frac{q_2}{q_3} 2\hat{z}_1^2 \hat{z}_3 - D_{13} \frac{q_2}{q_3} 3\hat{z}_1^2 \hat{z}_3^2 - C_3 \frac{q_2}{q_3} \hat{z}_1 \hat{z}_3^2 - C_1 \frac{q_2}{q_1} \hat{z}_1^2 \hat{z}_3 \qquad \text{(Eq. 28)}$$

Note that f_2^0 does not appear in Eq. 28; it drops out since $\hat{z}_1 + \hat{z}_3 = 1$. If we substitute Eq. 28 into Eq. 15, we obtain the final equation:

$$\ln x_{2,m}^s = \hat{z}_1 \ln x_{2,1}^s + \hat{z}_3 \ln x_{2,3}^s - A_{1\text{-}3} \hat{z}_1 \hat{z}_3 (2\hat{z}_1 - 1) \frac{q_2}{q_1}$$

$$+ A_{3\text{-}1} 2\hat{z}_1^2 \hat{z}_3 \frac{q_2}{q_3} + D_{13} 3\hat{z}_1^2 \hat{z}_3^2 \frac{q_2}{q_3} + C_3 \hat{z}_1 \hat{z}_3^2 \frac{q_2}{q_3} + C_1 \hat{z}_1^2 \hat{z}_3 \frac{q_2}{q_1} \qquad \text{(Eq. 29)}$$

The above equation will be referred to as the reduced four-suffix solubility equation since it was derived from the four-suffix excess free energy expression. Other solubility equations to be derived later will be referred to in a similar manner.

Equation 29 expresses the mole fraction solubility $x_{2,m}^s$ of the solute in the mixed solvent in terms of (a) the pure solvent solubilities given by the first two terms, (b) contributions from solvent–solvent interactions given by the next three terms, and (c) contributions from the interaction between the solute and the solvent mixture described by the last two terms. The data needed for the estimation of the solubility in Eq. 29 are:

1. The pure solvent solubilities $x_{2,1}^s$ and $x_{2,3}^s$.
2. The binary solvent data (usually vapor pressure as a function of composition) from which solvent–solvent interaction constants $A_{1\text{-}3}$, $A_{3\text{-}1}$, and D_{13} may be obtained,
3. The molar volumes of the pure compounds.
4. Two ternary experimental points (obtained from the solubility of the solute in the solvent mixture at two different compositions) from which C_1 and C_3 are calculated. It is important to note that the solvent–solvent interaction constants, once obtained at a particular temperature are fixed for the solvent system at that temperature.

Corresponding two- and three-suffix equations may also be written for the system described by Eq. 23. The three-suffix equation is obtained by setting the D values equal to zero or, in other words, by neglecting four-body interactions. If this is done, then:

$$C_2 = C_1 \frac{q_2}{q_1}, \quad C_3 = C_1 \frac{q_3}{q_1} \qquad \text{(Eq. 30)}$$

and Eq. 23 becomes:

$$\ln \gamma_{2,m} = z_3^2 \left[A_{2\text{-}3} + 2z_2 \left(A_{3\text{-}2} \frac{q_2}{q_3} - A_{2\text{-}3} \right) \right]$$

$$+ z_1^2 \left[A_{2\text{-}1} + 2z_2 \left(A_{1\text{-}2} \frac{q_2}{q_1} - A_{2\text{-}1} \right) \right]$$

$$+ z_1 z_3 \left[A_{3\text{-}2} \frac{q_2}{q_3} + A_{2\text{-}1} - A_{1\text{-}3} \frac{q_2}{q_1} + 2z_2 \left(A_{1\text{-}2} \frac{q_2}{q_1} - A_{2\text{-}1} \right) \right]$$

$$+ 2z_1 \left(A_{1\text{-}3} \frac{q_2}{q_1} - A_{3\text{-}1} \frac{q_2}{q_3} \right) - C_2 (1 - 2z_2) \right] \qquad \text{(Eq. 31)}$$

where $C_2 = q_2[3a_{112} + 3a_{133} + 3a_{223} - 6a_{123}]$ and the A values are as defined in Eqs. 22a–f without four-body interaction terms. Following the same treatment given in Eqs. 23–29, we arrive at the reduced three-suffix solubility equation:

$$\ln x_{2,m}^s = \hat{z}_1 \ln x_{2,1}^s + \hat{z}_3 \ln x_{2,3}^s - A_{1\text{-}3} \hat{z}_1 \hat{z}_3 (2\hat{z}_1 - 1) \frac{q_2}{q_1}$$

$$+ A_{3\text{-}1} 2\hat{z}_1^2 \hat{z}_3 \frac{q_2}{q_3} + C_2 \hat{z}_1 \hat{z}_3 \qquad \text{(Eq. 32)}$$

Equation 32 shows that only one constant, C_2, needs to be obtained from ternary data and, theoretically, only one point in the ternary system is needed.

The two-suffix equation is obtained by neglecting three- and four-body interactions. Then, Eq. 23 becomes:

$$\ln \gamma_{2,m} = z_3^2 A_{2\text{-}3} + z_1^2 A_{2\text{-}1} + z_1 z_3 \left(A_{2\text{-}3} + A_{2\text{-}1} - A_{1\text{-}3} \frac{A_{2\text{-}1}}{A_{1\text{-}2}} \right) \qquad \text{(Eq. 33)}$$

where

$$A_{1\text{-}2} = 2a_{12} q_1, \quad A_{2\text{-}1} = 2a_{12} q_2 \qquad \text{(Eq. 34a)}$$

$$A_{1\text{-}3} = 2a_{13} q_1, \quad A_{3\text{-}1} = 2a_{13} q_3 \qquad \text{(Eq. 34b)}$$

$$A_{2\text{-}3} = 2a_{23} q_2, \quad A_{3\text{-}2} = 2a_{23} q_3 \qquad \text{(Eq. 34c)}$$

Again, following the same treatment given in Eqs. 23–29 (excepting that it is not necessary to make the van Laar-type approximation mentioned earlier because it is already implied in Eq. 34), Eq. 33 leads to the reduced two-suffix solubility equation:

Table I—Summary of Equations [a]

Starting Equation	Differential	Simplifications	Final Equation
	Four-Suffix Excess Free Energy Model		
$[g^E/(RT)(x_1q_1 + x_2q_2 + x_3q_3)] = 2a_{12}z_1z_2$ $+ 2a_{13}z_1z_3 + 2a_{23}z_2z_3$ $+ 3a_{112}z_1^2z_2 + 3a_{122}z_1z_2^2$ $+ 3a_{113}z_1^2z_3 + 3a_{133}z_1z_3^2$ $+ 3a_{223}z_2^2z_3 + 3a_{233}z_2z_3^2$ $+ 6a_{123}z_1z_2z_3 + 4a_{1112}z_1^3z_2$ $+ 6a_{1122}z_1^2z_2^2 + 4a_{1222}z_1z_2^3$ $+ 4a_{1113}z_1^3z_3 + 6a_{1133}z_1^2z_3^2$ $+ 4a_{1333}z_1z_3^3 + 4a_{2223}z_2^3z_3$ $+ 6a_{2233}z_2^2z_3^2 + 4a_{2333}z_2z_3^3$ $+ 12a_{1123}z_1^2z_2z_3 + 12a_{1223}z_1z_2^2z_3$ $+ 12a_{1233}z_1z_2z_3^2$	$\ln \gamma_{2,m} = [\delta n_T g^E/RT/\delta n_2]_{T,P,n_1,n_3}$ $= z_3^2[A_{2-3} + 2z_2(A_{3-2}q_2/q_3 - A_{2-3}$ $- D_{23}) + 3z_2^2D_{23}] + z_1^2[A_{2-1}$ $+ 2z_2(A_{1-2}q_2/q_1 - A_{2-1} - D_{12}q_2/q_1)$ $+ 3z_2^2D_{12}q_2/q_1] + z_1z_3[A_{3-2}q_2/q_3$ $+ A_{2-1} - A_{1-3}q_2/q_1 + 2z_2(A_{1-2}q_2/q_1$ $- A_{2-1}) + 2z_1(A_{1-3}q_2/q_1$ $- A_{3-1}q_2/q_3) - 3z_1z_3D_{13}q_2/q_3$ $- C_2z_2(2 - 3z_2) - C_3q_2/q_3z_3(1$ $- 3z_2) - C_1q_2/q_1z_1(1 - 3z_2)]$	1. $3a_{112} = 3a_{122}$ $3a_{233} = 3a_{223}$ $4a_{1112} = 4a_{1222}$ $4a_{2333} = 4a_{2223}$ Then, $A_{1-2}/A_{2-1} = q_1/q_2$ and $A_{2-3}/A_{3-2} = q_2/q_3$ 2. Let $x_2 \to 0$ 3. $\lim_{x_2\to 0} \ln \gamma_{2,1}$ $= \ln (H_{2,1}/f_2^0) = A_{2-1}$ and $\lim_{x_2\to 0} \ln \gamma_{2,3}$ $= \ln (H_{2,3}/f_2^0)$ $= A_{2-3}$	$\ln x_{2,m}^s = \hat{z}_1 \ln x_{2,1}^2$ $+ \hat{z}_3 \ln x_{2,3}^s - A_{1-3}\hat{z}_1\hat{z}_3$ $(2\hat{z}_1 - 1)q_2/q_1 +$ $A_{3-1}2\hat{z}_1^2\hat{z}_3q_2/q_3$ $+ D_{13}3\hat{z}_1^2\hat{z}_3^2q_2/q_3$ $+ C_3\hat{z}_1^2\hat{z}_3^2q_2/q_3$ $+ C_1\hat{z}_1^2\hat{z}_3q_2/q_1$
	Three-Suffix Excess Free Energy Model		
Four-suffix equation with: $a_{1112} = a_{1122}$ $= a_{1222} = a_{1113}$ $= a_{1133} = a_{1333} = a_{2223}$ $= a_{2233} = a_{2333}$ $= a_{1123} = a_{1223} = a_{1233} = 0$	$\ln \gamma_{2,m}$ $= z_3[A_{2-3} + 2z_2(A_{3-2}q_2/q_3 - A_{2-3})]$ $+ z_1^2[A_{2-1} + 2z_2(A_{1-2}q_2/q_1 - A_{2-1})]$ $+ z_1z_3[A_{3-2}q_2/q_3 + A_{2-1} - A_{1-3}q_2/q_1$ $+ 2z_2(A_{1-2}q_2/q_1 - A_{2-1}) + 2z_1(A_{1-3}$ $q_2/q_1 - A_{3-1}q_2/q_3) - C_2(1 - 2z_2)]$	1. $3a_{112} = 3a_{122}$ $3a_{233} = 3a_{223}$ Then, $A_{1-2}/A_{2-1} = q_1/q_2$ and $A_{2-3}/A_{3-2} = q_2/q_3$ 2. Simplifications 2 and 3 from the four-suffix model	$\ln x_{2,m}^s = \hat{z}_1 \ln x_{2,1}^s$ $+ \hat{z}_3 \ln x_{2,3}^s - A_{1-3}$ $\hat{z}_1\hat{z}_3(2\hat{z}_1 - 1)q_2/q_1$ $+ A_{3-1}2\hat{z}_1^2\hat{z}_3q_2/q_3$ $+ C_2\hat{z}_1\hat{z}_3$
	Two-Suffix Excess Free Energy Model		
Four-suffix equation with: $a_{112} = a_{122} = a_{113} = a_{133} = a_{223}$ $= a_{233} = a_{123} = \ldots = 0$	$z_3^2A_{2-3} + z_1^2A_{2-1} + z_1z_3(A_{2-3}$ $+ A_{2-1} - A_{1-3}A_{2-1}/A_{1-2})]$	1. Simplifications 2 and 3 from the four-suffix model	$\ln x_{2,m}^s = \hat{z}_1 \ln x_{2,1}^s$ $+ \hat{z}_3 \ln x_{2,3}^s + A_{1-3}\hat{z}_1\hat{z}_3$ q_2/q_1

[a] The symbols are defined in the text. Note that the A values for the two-suffix equation are defined differently than those for the three- and four-suffix equations.

$$\ln x_{2,m}^s = \hat{z}_1 \ln x_{2,1}^s + \hat{z}_3 \ln x_{2,3}^s + A_{1-3}\hat{z}_1\hat{z}_3\frac{q_2}{q_1} \quad \text{(Eq. 35)}$$

Examination of Eq. 35 shows that unlike the reduced three- and four-suffix solubility equations, it contains no constant to be estimated from ternary data. This, combined with the fact that the reduced three-suffix solubility equation with its fewer parameters (compared with the corresponding four-suffix equation) satisfactorily describes the data used (6), leads us to choose Eq. 32 as our working equation for a three-component system.

Four-Component Systems—For a four-component system, the three-suffix excess free energy expression is given by:

$$\frac{g^E}{RT\left(x_1 + x_2\frac{q_2}{q_1} + x_3\frac{q_3}{q_1} + x_4\frac{q_4}{q_1}\right)} = q_1(2a_{12}z_1z_2 + 2a_{13}z_1z_3$$
$$+ 2a_{14}z_1z_4 + 2a_{23}z_2z_3 + 2a_{34}z_3z_4 + 2a_{24}z_2z_4 + 3a_{112}z_1^2z_2$$
$$+ 3a_{122}z_1z_2^2 + 3a_{113}z_1^2z_3 + 3a_{133}z_1z_3^2 + 3a_{114}z_1^2z_4$$
$$+ 3a_{144}z_1z_4^2 + 3a_{223}z_2^2z_3 + 3a_{233}z_2z_3^2 + 3a_{224}z_2^2z_4$$
$$+ 3a_{244}z_2z_4^2 + 3a_{334}z_3^2z_4 + 3a_{344}z_3z_4^2 + 6a_{123}z_1z_2z_3$$
$$+ 6a_{124}z_1z_2z_4 + 6a_{234}z_2z_3z_4 + 6a_{134}z_1z_3z_4$$
$$+ 24a_{1234}z_1z_2z_3z_4) \quad \text{(Eq. 36)}$$

With the following definitions:

$$A_{1-4} = q_1(2a_{14} + 3a_{144}) \quad \text{(Eq. 37a)}$$

$$A_{4-1} = q_4(2a_{14} + 3a_{114}) \quad \text{(Eq. 37b)}$$

$$A_{3-4} = q_3(2a_{34} + 3a_{344}) \quad \text{(Eq. 37c)}$$

$$A_{4-3} = q_4(2a_{34} + 3a_{334}) \quad \text{(Eq. 37d)}$$

$$A_{2-4} = q_2(2a_{24} + 3a_{244}) \quad \text{(Eq. 37e)}$$

$$A_{4-2} = q_4(2a_{24} + 3a_{224}) \quad \text{(Eq. 37f)}$$

$$G_{123} = 3a_{112} + 3a_{133} + 3a_{223} - 6a_{123} \quad \text{(Eq. 37g)}$$

$$G_{234} = 3a_{223} + 3a_{334} + 3a_{244} - 6a_{234} \quad \text{(Eq. 37h)}$$

$$G_{134} = 3a_{144} + 3a_{113} + 3a_{334} - 6a_{134} \quad \text{(Eq. 37i)}$$

$$G_{124} = 3a_{112} + 3a_{224} + 3a_{144} - 6a_{124} \quad \text{(Eq. 37j)}$$

$$K = 24a_{1234} \quad \text{(Eq. 37k)}$$

Eq. 36 becomes:

$$\frac{g^E}{RT\left(x_1 + x_2\frac{q_2}{q_1} + x_3\frac{q_3}{q_1} + x_4\frac{q_4}{q_1}\right)} = z_1z_2\left(A_{2-1}z_1\frac{q_1}{q_2} + A_{1-2}z_2\right)$$
$$+ z_1z_3\left(A_{3-1}z_1\frac{q_1}{q_3} + A_{1-3}z_3\right) + z_1z_4\left(A_{4-1}z_1\frac{q_1}{q_4} + A_{1-4}z_4\right)$$
$$+ z_2z_3\left(A_{3-2}z_2\frac{q_1}{q_3} + A_{2-3}z_3\frac{q_1}{q_2}\right) + z_2z_4\left(A_{4-2}\frac{q_1}{q_4} + A_{2-4}z_4\frac{q_1}{q_2}\right)$$
$$+ z_3z_4\left(A_{4-3}z_3\frac{q_1}{q_4} + A_{3-4}z_4\frac{q_1}{q_3}\right) + z_1z_2z_3\left(A_{2-1}\frac{q_1}{q_2} + A_{3-2}\frac{q_1}{q_3} + A_{1-3}\right.$$
$$\left. - G_{123}q_1\right) + z_1z_2z_4\left(A_{2-1}\frac{q_1}{q_2} + A_{4-2}\frac{q_1}{q_4} + A_{1-4} - G_{124}q_1\right)$$
$$+ z_2z_3z_4\left(A_{3-2}\frac{q_1}{q_3} + A_{4-3}\frac{q_1}{q_4} + A_{2-4}\frac{q_1}{q_2} - G_{234}q_1\right) + z_1z_3z_4\left(A_{1-4}\right.$$
$$\left. + A_{3-1}\frac{q_1}{q_3} + A_{4-3}\frac{q_1}{q_4} - G_{134}q_1\right) + Kz_1z_2z_3z_4q_1 \quad \text{(Eq. 38)}$$

where A_{1-2}, A_{2-1}, A_{1-3}, A_{3-1}, A_{2-3}, and A_{3-2} are as defined in Eqs. 22a–f without the four-body interaction terms. If we consider this system as a solute (denoted by subscript 2) in a mixture of three solvents (denoted by subscripts 1, 3, and 4), then an approach similar to that given in Eqs. 23 to 29 yields the reduced three-suffix solubility equation for a four-component system:

$$\ln x_{2,m}^s = \hat{z}_1 \ln x_{2,1}^s + \hat{z}_3 \ln x_{2,3}^s + \hat{z}_4 \ln x_{2,4}^s - A_{1-3}\hat{z}_1z_3(2\hat{z}_1 + 2\hat{z}_4 - 1)\frac{q_2}{q_1}$$
$$+ A_{3-1}2\hat{z}_1\hat{z}_3(\hat{z}_1 + \hat{z}_4)\frac{q_2}{q_3} - A_{1-4}\hat{z}_1\hat{z}_4(2\hat{z}_1 - 1)\frac{q_2}{q_1} + A_{4-1}2\hat{z}_1^2\hat{z}_4\frac{q_2}{q_4}$$
$$- A_{4-3}\hat{z}_3\hat{z}_4(2\hat{z}_4 - 1)\frac{q_2}{q_4} + A_{3-4}2\hat{z}_3\hat{z}_4^2\frac{q_2}{q_3} + G_{123}q_2\hat{z}_1\hat{z}_3$$

$$+ G_{124}q_2\hat{z}_1\hat{z}_4 + G_{234}q_2\hat{z}_3\hat{z}_4 - G_{134}q_2\hat{z}_1\hat{z}_3\hat{z}_4 - Kq_2\hat{z}_1\hat{z}_3\hat{z}_4 \quad \text{(Eq. 39)}$$

Equation 39 shows that only one constant, K, need be obtained from a quaternary system. The A values and G_{134} are obtained from solvent vapor pressure composition data. If we let $\hat{z}_4$ equal zero, Eq. 39 reduces to Eq. 32 with $C_2 = G_{123}q_2$. Thus G_{123}, G_{124}, and G_{234} are estimated from solubility data at points where solvent 4, solvent 3, and solvent 1, respectively, are absent. They are, therefore, not new parameters in the same sense as K. It is important to note that even with the more complex four-component system, we need to estimate only one quaternary constant to describe the system. It is clear that if Eq. 39 is used for predictive purposes (for systems in which it is valid) the experimental labor saved is substantial. The same consideration holds to a lesser degree for Eq. 32.

DISCUSSION

The reduced three-suffix solubility equation for a solute in a mixture of two solvents (Eq. 32) describes solubility in terms of pure solvent solubilities and contributions from solvent–solvent and solute–binary solvent interactions. The advantages of this approach over others are its generality and flexibility. Once the solvent–solvent interaction constants are obtained, they are fixed for that solvent system (provided one operates at or close to the temperature at which they were determined, usually 25°C). Thus, the only constant that needs to be determined to completely describe the solubility (assuming that the solubility of the solute in each solvent is known) is the ternary constant C_2. For a solute in a mixture of three solvents (four-component system), only one quaternary constant (K in Eq. 39) is needed to describe the solubility. The ternary constants G_{124}, G_{234}, and G_{123} are obtained in practice in the absence of solvents 3, 1, and 4, respectively. The essential aspects of this treatment are presented in Table I, including the n-suffix excess free energy model for the number of components involved and assumptions which simplify the solute–solvent interactions while keeping the solvent–solvent interaction terms as detailed as necessary.

REFERENCES

(1) M. A. Moustafa, A. M. Molokhia, and M. W. Gouda, *J. Pharm. Sci.*, **70**, 1172 (1981).
(2) A. Martin, J. Newburger, and A. Adjei, *J. Pharm. Sci.*, **69**, 487 (1980).
(3) A. Adjei, J. Newburger, and A. Martin, *J. Pharm. Sci.*, **69**, 659 (1980).
(4) A. Martin, A. N. Paruta, and A. Adjei, *J. Pharm. Sci.*, **70**, 1115 (1981).
(5) A. Martin, P. L. Wu, A. Adjei, A. Beerbower, and J. M. Prausnitz, *J. Pharm. Sci.*, **70**, 1260 (1981).
(6) N. A. Williams and G. L. Amidon, **73**, 14 (1984).
(7) N. A. Williams and G. L. Amidon, **73**, 18 (1984).
(8) J. M. Prausnitz, "Molecular Thermodynamics of Fluid-Phase Equilibria," Prentice-Hall, Englewood Cliffs, N.J., 1969, chap. 6.
(9) I. Prigogine and R. Defay, "Chemical Thermodynamics," Longmans, London, 1954, chap. 21.
(10) K. Wohl, *Trans. Am. Inst. Chem. Eng.*, **42**, 215 (1946).