Pharmaceutical Research, Vol. 5, No. 3, 1988

Report

The Estimation of Solubility in Binary Solvents: Application of the Reduced 3-Suffix Solubility Equation to Ethanol–Water Mixtures

N. Adeyinka Williams[1,2] and Gordon L. Amidon[1]

Received July 21, 1987; accepted November 4, 1987

The reduced 3-suffix solubility equation (R3SSE) is applied to the characterization of solubility in the ethanol–water system. The data needed are the solubility of the compound in each of the pure solvents and at one ethanol–water composition. This composition has been estimated from solubility data to be 0.56 volume fraction of ethanol. The solubility obtained at this volume fraction is used to estimate the ternary solute–solvent interaction constant, C_2. The R3SSE, with the C_2 thus obtained, predicts the mixed solvent solubilities of the compounds tested, as accurately as that obtained from several volume fractions. The superiority of the R3SSE over two related equations—a simple second-degree polynomial equation and a simplified form of the R3SSE which neglects contributions to solubility from the solvent mixture—is also demonstrated for a number of solutes.

KEY WORDS: reduced 3-suffix solubility equation; ethanol–water; mixed solvents.

INTRODUCTION

Studies involving the modeling and prediction of solubility continue to be of interest (1–6), not only because they improve our knowledge of solution thermodynamics but also because they provide a tool for the prediction of solubilities in solvent mixtures with a minimum of solubility determinations.

The reduced 3-suffix solubility equation (R3SSE) was developed to predict the solubility of nonpolar organic compounds in solvent mixtures (4). The equation was shown to be adequate for describing the solubility profiles of the compounds tested in ethanol–water, propylene glycol–water, and ethanol–propylene glycol–water systems (5,6). For a binary solvent mixture, the experimental data needed are the solubility of the compound in each of the pure solvents and the solubility in one mixture of the solvents. The latter is needed to estimate the solute–solvent interaction constant C_2. Previously, C_2 was estimated from experimental points (ranging from 4 to 13 in number) determined over the whole solvent composition range (5). However, the more convenient method is to estimate C_2 from one experimental solubility determined from a cosolvent volume (or weight) fraction of somewhere between zero and one.

Since the solvent–solvent interaction parameters depend only on temperature, once obtained for a mixed solvent system, they are fixed. Therefore, at a given temperature, the R3SSE may be reduced to a one-parameter equation with C_2 as the only undetermined parameter.

In this paper, we show how C_2 may be estimated from a one-parameter equation as would be done in practice. We also show that the C_2 thus obtained is as effective as that obtained from several points in the solvent composition range. Finally, we examine some equations related to the R3SSE and evaluate their effectiveness compared to the R3SSE.

METHODS

The R3SSE for Ethanol–Water at 25°C

For a binary solvent system such as ethanol and water (represented by subscripts 1 and 3, respectively), the R3SSE is

$$\begin{aligned}\ln X^{s}_{2,m} = {} & \ln X^{s}_{2,3} + z_1 \ln(X^{s}_{2,1}/X^{s}_{2,3}) - A_{1-3}\, z_1\, z_3 \\ & (2\, z_1 - 1)\, q_2/q_1 + A_{3-1}\, 2\, z_1^2\, z_3\, q_2/q_3 \\ & + C_2\, z_1\, z_3 \ldots \qquad (1)\end{aligned}$$

where $X^{s}_{2,m}$ is the solubility of the solute in the mixed solvent; $X^{s}_{2,1}$ is the solubility in pure solvent 1 (ethanol); $X^{s}_{2,3}$ is the solubility in pure solvent 3 (water); A_{1-3} and A_{3-1} are the constants characteristic of the interaction between solvent 1 and solvent 3; q_1, q_2, and q_3 are the molar volumes of solvent 1, the solute and solvent 3, respectively; and z_1 and z_3 are the volume fractions of solvents 1 and 3, respectively. In previous papers (4–6), z_1 and z_3 were written, where appropriate, as $\hat{z}_1$ and $\hat{z}_3$ to indicate that they were solute-free volume fractions but the caret is omitted here for simplicity.

For the ethanol–water system at 25°C,

$A_{1-3} = 1.216$, $q_1 = 58.68$ cm^3/mol
$A_{3-1} = 0.9093$, $q_3 = 18.07$ cm^3/mol

---

[1] College of Pharmacy, The University of Michigan, Ann Arbor, Michigan 48109-1065.

[2] To whom correspondence should be addressed.

Noting that $z_3 = 1 - z_1$ and substituting the values given for the constants above, Eq. (1) may be rearranged to give

$$\ln X_{2,m}^{s} = \ln X_{2,3}^{s} + z_1 \ln(X_{2,1}^{s}/X_{2,3}^{s}) + \{[0.080 - 0.059\,(1 - z_1)]q_2 + C_2\}\,(1 - z_1)z_1 \qquad (2)$$

Thus for ethanol–water at 25°C, the solubility is given by Eq. (2). The molar volume of the solute, q_2, may be determined by measuring the density and multiplying its reciprocal by the molecular weight or it may be estimated. Therefore, only the constant C_2 needs to be determined in order to generate the solubility profile of the solute.

One-Point Estimation of C_2

In previous papers (5,6) C_2 was evaluated from the equation

$$C_2 = \frac{(\ln X_{2,m}^{s})_{\mathrm{exp}} - (\ln X_{2,m}^{s})_{\mathrm{calc}}}{z_1\,(1 - z_1)} \qquad (3)$$

where $(\ln X_{2,m}^{s})_{\mathrm{exp}}$ is the natural logarithm of the experimental solubility at the given cosolvent volume fraction, and $(\ln X_{2,m}^{s})_{\mathrm{calc}}$ is the natural logarithm of the solubility calculated from Eq. (2) with $C_2 = 0$. In this manner, C_2, for each solute, was evaluated from as many experimental solubility points as available. Then the best value of C_2 was obtained by a least-squares fit of the data for each solute concerned. In practice, however, where the aim is to generate a solubility profile with as little experimental labor as possible, several experimental determinations to evaluate C_2 obviously defeat the purpose. In this paper we show that the C_2 value determined from one solubility measurement at a point in the solvent composition range gives a solubility profile which compares satisfactorily (for the compounds tested) with that obtained by the previous method. For convenience, the C_2 obtained from several points is designated $C_{2,\mathrm{avg}}$ and that obtained from one point is designated $C_{2,0}$.

Estimation of $C_{2,0}$

$C_{2,0}$ was estimated as follows: for each solute, a graph of C_2 calculated from Eq. (3) was plotted against z_1 using experimental points. The volume fraction at which the curve intersected the line given by the value of $C_{2,\mathrm{avg}}$ was noted and the average of such volume fractions, obtained for nine solutes,[3] was 0.56, with a standard deviation of 0.09. Next, $C_{2,0}$ was estimated for each of the compounds tested in this study at a cosolvent volume fraction of 0.56 and used in generating the solubility profiles for the compounds (see Table I).

Evaluation of Equations Related to the R3SSE

It was also of interest to evaluate the usefulness of two other equations—a simple second-degree polynomial and a simplified form of the R3SSE—in describing the solubility of these compounds.

The simple second-degree polynomial is an empirical equation proposed by Yalkowsky and Roseman (7) for describing solubility in mixed solvents and may be written as

$$\ln X_{2,m}^{s} = \ln X_{2,3}^{s} + z_1\,\theta_1 + z_1^2\theta_2 \ldots \qquad (4)$$

where $X_{2,m}^{s}$, $X_{2,3}^{s}$, and z_1 are as defined previously and θ_1 and θ_2 are constants (note that the subscript in θ_2 does not refer to the solute). The simplified R3SSE is obtained by setting the solvent–solvent constants A_{1-3} and A_{3-1} equal to zero. Thus, the equation reduces to

$$\ln X_{2,m}^{s} = \ln X_{2,3}^{s} + z_1 \ln(X_{2,1}^{s}/X_{2,3}^{s}) + \beta(1 - z_1)z_1 \ldots \qquad (5)$$

From Eqs. (2) and (5), it is seen that

$$\beta = [0.080 - 0.059\,(1 - z_1)]q_2 + C_2$$

Equation (5), in effect, ignores the volume-fraction dependence of the solvent–solvent interactions.

It should be noted that while Eq. (5) reduces to $\ln X_{2,1}^{s}$ at $z_1 = 1$ as it should, Eq. (4) does not. It is clear, therefore, that even if Eq. (4) performs well in describing solubility between the two end points of the solvent composition, it

---

[3] Phenobarbital, phenylsalicylate, stearic acid, acetanilide, DL-valine, DL-alanine, glycine, barbital, and *o*-nitrophenol were used.

Table I. A Comparison of the Various Equations in Terms of Average Errors in the Predicted Solubility

	Average relative error (%)[a]					
Compound	R3SSE with $C_{2,\mathrm{avg}}$	R3SSE with $C_{2,0}$	Second-degree polynomial equation [Eq. (4)]	Eq. (5)	No. of experimental points used[b]	Ref. No. for solubility data
O-Nitrophenol	6.8	8.6	21.2	19.9	9	8
Phenobarbital	4.7	5.3	24.8	31.5	15	9
Glycine	19.8	16.9	25.5	99.8	7	10
Phenylsalicylate	25.3	29.1	19.3	18.0	9	11
Acetanilide	7.3	7.1	18.6	23.7	13	11
Caffeine	10.8	10.9	55.9	41.1	7	12[c]
Salicylic acid	17.1	18.1	55.6	34.8	6	12

[a] Calculated as $\dfrac{|(X_{2,m}^{s})_{\mathrm{exp}} - (X_{2,m}^{s})_{\mathrm{calc}}|}{(X_{2,m}^{s})_{\mathrm{exp}}} \times 100.$

[b] Includes both end points.

[c] The solubility of caffeine in ethanol was obtained from Ref. 13.

will fail as z_1 approaches 1. In other words, the form of Eq. (4) is actually inconsistent with the known characteristics of the system.

RESULTS AND DISCUSSION

Figures 1 and 2 show solubility profiles of caffeine and salicylic acid generated using the R3SSE with $C_{2,\mathrm{avg}}$ and the R3SSE with $C_{2,0}$ as the ternary constants. Solubility profiles for the rest of the compounds used have been shown in a previous paper (5). For such compounds (including the two named above), the average relative errors of their predicted solubilities using the two kinds of C_2 are shown in Table I. It is clear from Table I and Figs. 1 and 2 that the R3SSE with $C_{2,0}$ is quite comparable in accuracy to the R3SSE with $C_{2,\mathrm{avg}}$. This result is significant (at least for the ethanol–water system) because it indicates that—apart from the pure solvent solubilities—only one experimental solubility (at a cosolvent volume fraction of 0.56) needs to be obtained in order to generate a complete curve of solubility versus solvent composition. As pointed out previously (4), this statement is correct in theory but it could not be said, then, that a C_2 obtained from one point would be good enough, nor was it clear where, in the solvent composition range, one should estimate it. Since we have not attempted this kind of analysis for other binary solvent mixtures, we cannot say if C_2 should always be estimated at a cosolvent volume fraction of 0.56 regardless of the solvent system. With the ethanol–water system, the calculated solubility without the C_2 term often shows its greatest deviation from the measured solubility at an ethanol volume fraction of about 0.4 to 0.6. We expect other binary solvent systems to show somewhat similar characteristics. Therefore, for such systems, C_2 may be estimated at a cosolvent volume fraction of 0.56 if there are no data from which to obtain the "optimum" volume fraction.

Table I also shows the results of a comparison between the R3SSE with $C_{2,0}$ and the two related equations discussed earlier. The end points, i.e., the solubilities at $z_1 = 0$ and $z_1 = 1$, were included in the averages because predicting the experimental solubility at each end point is also a measure of a model's effectiveness and consistence with the characteristics of the system. From Table I, it is seen that the R3SSE with $C_{2,0}$ is more accurate than Eqs. (4) and (5) in predicting solubility for all the solutes tested except for phenylsalicylate, where it is less accurate by about 10% relative to the two equations. On the other hand, Eqs. (4) and (5) perform much worse in predicting solubility for the other compounds in the table. Their poor performance justifies the need, it appears, for a more accurate equation such as the R3SSE which takes into account solvent–solvent interactions and also reduces to the pure solvent solubilities at each end point.

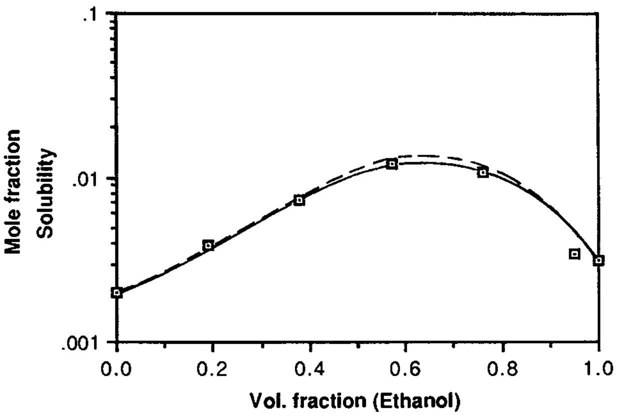

Fig. 1. Solubility of caffeine in ethanol–water mixtures at 25°C. (——) R3SSE with $C_{2,\mathrm{avg}}$; (----) R3SSE with $C_{2,0}$; (□) experimental points.

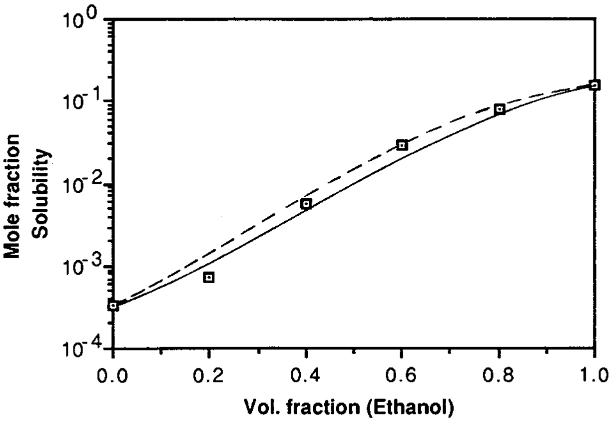

Fig. 2. Solubility of salicylic acid in ethanol–water mixtures at 25°C. (——) R3SSE with $C_{2,\mathrm{avg}}$; (----) R3SSE with $C_{2,0}$; (□) experimental points.

In summary, the R3SSE has been shown to be the preferred one-parameter equation with which to estimate solubility in a binary solvent system. Although slightly more complex in form, it is more accurate than other one-parameter equations.

REFERENCES

1. A. Martin, J. Newburger, and A. Adjei. *J. Pharm. Sci.* 69:487–491 (1980).
2. W. E. Acree, Jr., and J. H. Rytting. *J. Pharm. Sci.* 72:292–296 (1983).
3. J. T. Rubino and S. H. Yalkowsky. *Pharm. Res.* 4:220–230 (1987).
4. N. A. Williams and G. L. Amidon. *J. Pharm. Sci.* 73:10–13 (1984).
5. N. A. Williams and G. L. Amidon. *J. Pharm. Sci.* 73:14–18 (1984).
6. N. A. Williams and G. L. Amidon. *J. Pharm. Sci.* 73:18–23 (1984).
7. S. H. Yalkowsky and T. J. Roseman. In S. H. Yalkowsky (ed.), *Techniques of Solubilization of Drugs,* Marcel Dekker, New York, 1981, Chap. 3.
8. J. C. Duff and E. J. Bills. *J. Chem. Soc.* 1930:1331–1338 (1930).
9. T. L. Breon and A. N. Paruta. *J. Pharm. Sci.* 59:1306–1313 (1970).
10. J. P. Greenstein and M. Winitz. *Chemistry of the Amino Acids, Vol. 1,* Wiley, New York, 1961, p. 547.
11. H. Stephen and T. Stephen. *Solubilities of Inorganic and Organic Compounds, Vol. 2, Part 2,* Macmillan, New York, 1964.
12. S. Yalkowsky. Personal communication.
13. A. Seidell. *Solubilities of Organic Compounds, Vol. 2,* Van Nostrand, New York, 1941.