ELSEVIER

International Journal of Pharmaceutics 127 (1996) 27–30

international journal of pharmaceutics

Comments concerning 'Model for solubility estimation in mixed solvent systems'

William E. Acree, Jr.

Department of Chemistry, University of North Texas, Denton, TX 76203-0068, USA

Received 30 November 1994; revised 25 January 1995; accepted 9 May 1995

In a recent paper appearing in this journal, Barzegar-Jalali and Hanaee (1994) proposed a double log-log model for the linearization and accurate mathematical representation of experimental solubility data for mixed aqueous-organic solvent systems. Depending upon initial solvent composition, solute (component A) solubility was expressed either as

$$\begin{aligned}\ln\{\ln[x_A^{sat}/(x_A^{sat})_W]\} &= \ln\{\ln[(x_A^{sat})_{0.5}/(x_A^{sat})_W]\} \\ &\quad + B\ln(\phi_C^\circ/\phi_W^\circ) \\ &= \text{intercept} \\ &\quad + \text{slope}\ \ln(\phi_C^\circ/\phi_W^\circ) \qquad (1)\end{aligned}$$

when

$$0 < \phi_C^\circ \leq 0.5,$$

or

$$\begin{aligned}\ln\{\ln[(x_A^{sat})_C/x_A^{sat}]\} &= \ln\{\ln[(x_A^{sat})_C/(x_A^{sat})_{0.5}]\} \\ &\quad + b\ln(\phi_W^\circ/0.5) \\ &= \text{intercept} \\ &\quad + \text{slope}\ \ln(\phi_W^\circ/0.5) \qquad (2)\end{aligned}$$

when $0 < \phi_W^\circ \leq 0.5$, where x_A^{sat}, $(x_A^{sat})_W$, $(x_A^{sat})_{0.5}$ and $(x_A^{sat})_C$ are the drug solubilities in the mixed solvent system, water, a system containing 0.5 fraction of the organic solvent (and/or water) and neat organic solvent, respectively. Compositions $\phi_C^\circ\}$ and ϕ_W° denote the volume fractions of the organic solvent and water calculated as if the solute were not present. To demonstrate the applicability of Eqs. (1) and (2), the authors showed that the descriptive accuracy of their log-log model was superior to expressions derived previously by Williams and Amidon (1984) based upon an excess free energy model. In these computations the slopes (and possibly the intercepts) were treated as 'curve-fit' parameters determined via least squares regressional analysis, and comparisons were limited to eight alkyl *p*-hydroxybenzoates and alkyl *p*-amino-benzoates dissolved in aqueous-propylene glycol mixtures.

The purpose of this communication is not to criticize work of Barzegar-Jalali and Hanaee. Rather, I wish to address a much broader issue, namely the mathematical representation of isothermal solute solubility data measured in binary solvent mixtures. While linearized forms such as Eqs. (1) and (2) were once popular, today's computer technology and readily available commercial software packages allow researchers significantly greater freedom in reporting how the solubility varies with solvent composition. Recommendation of a particular mathematical form

SSDI 0378-5173(95)04109-N

for generalized curve-fitting of solubility data should not be dictated solely by computational simplicity. The equation's ability to accurately describe the measured solubility data and its applicability to a wide range of solute-solvent systems are far more important considerations.

Expressions for predicting the thermodynamic properties of ternary non-electrolyte systems have served as the point of departure for mathematical representation of experimental excess molar Gibbs energy, excess molar heat capacity, excess molar enthalpy and excess molar volume data. Differences between predicted and observed values are expressed as

$$(Z^{E}_{ABC})^{exp} - (Z^{E}_{ABC})^{calc} = x_A x_B x_C Q_{ABC} \quad (3)$$

with Q-functions of varying complexity. Acree and Zvaigzne (Acree and Zvaigzne, 1991) extended these ideas to isothermal solubility data. The authors suggested a Combined NIBS/Redlich-Kister mathematical representation

$$\ln x^{sat}_A = x^{\circ}_B \ln(x^{sat}_A)_B + x^{\circ}_C \ln(x^{sat}_A)_C + x^{\circ}_B x^{\circ}_C \sum_{i=0}^{N} S_i (x^{\circ}_B - x^{\circ}_C)^i \quad (4)$$

where B and C denote the two solvent components and the various S_i 'curve-fit' parameters can be evaluated via least squares analysis. Applicability of Eq. (4) has been previously established (Acree, 1994) using experimental solubility data for anthracene, pyrene and carbazole in a wide range of binary organic solvent mixtures. Eq. (4) is derivable from a realistical solution model, with the 'curve-fit' coefficients being related to solute-solute, solute-solvent and solvent-solvent energy differences (Acree, 1992). The double log-log model, on the other hand, is a strictly empirical equation developed solely for mathematical representation of experimental solubility data.

To provide a fair tête-à-tête comparison of the descriptive ability of the double log-log model versus Combined NIBS/Redlich Kister model is extremely difficult. The double log-log model contains separate equations for describing solubility data in organic solvent rich and in water rich binary solvent mixtures. Each separate equation has either one or two curve-fit coefficients, depending upon how the intercept is treated. Eq. (4), on the other hand, is used for the entire solvent composition range. With this idea in mind, I contend that the most impartial assessment of the relative descriptive abilities would be to allow Eq. (4) three parameters to help offset any superiority that might be gained as a result of the double log-log model's additional equation.

The data set used by Barzegar-Jalali and Hanaee is reported as mole fraction solubilities at nine specific propylene glycol volume fractions spanning the entire binary composition range. Rather than convert everything to mole fractions, the Combined NIBS/Redlich-Kister equation will be modified to

$$\ln x^{sat}_A = \phi^{\circ}_C \ln(x^{sat}_A)_C + \phi^{\circ}_W \ln(x^{sat}_A)_W + \phi^{\circ}_C \phi^{\circ}_W \sum_{i=0}^{N} S_i (\phi^{\circ}_C - \phi^{\circ}_W)^i \quad (5)$$

Readers may recognize the first two terms on the right-hand side of Eq. (5) as the log mole fraction solubility analog of the very popular log-linear solubility equation derived and verified by Yalkowsky and coworkers (Yalkowsky and Roseman, 1981; Gupta et al., 1991), or as Equation XV of the Nearly Ideal Binary Solvent (NIBS) model (Acree and Bertrand, 1977) with the solvent-solvent unmixing term set equal to zero. Table 1 compares the descriptive abilities of the double log-log model and Eq. (5). Summed squared percentage deviation is

$$\sum(\%Dev)^2 = \sum\{[[(x^{sat}_A)^{calc} - (x^{sat}_A)^{exp}]/(x^{sat}_A)^{exp}]\}^2 \quad (6)$$

taken as a measure of the 'goodness-of-fit' of each model. All entries pertaining to the double log-log model were taken directly from the work of Barzegar-Jalali and Hanaee. In the case of Eq. (5), both the two- and three-parameter forms were considered. To conserve space only the latter set of parameters are reported. Careful examination of Table 1 reveals that Eq. (5) is far superior to the double log-log model, particularly in the case of propyl p-hydroxybenzoate, butyl p-hydroxybenzoate and butyl p-aminobenzoate dissolved in aqueous-propylene glycol mixtures. These three

Table 1
Summarized comparison of the descriptive abilities of the double log-log model and Eq. (5)

Solute[a]	Double log-log model			Eq. (5)	
	Eq. 1[b]	Eq. 2[b]	$(\Sigma\ \%\ \mathrm{Dev})^2$	S_i^c	$(\Sigma\ \%\ \mathrm{Dev})^2$
MPHB	0.9400	1.3180	288.8	−0.049	318
	1.0787	1.0173		3.424	
				0.255	
EPHB	0.7744	1.2126	568.3	−0.591	620
	1.2054	1.2109		3.438	
				0.016	
PPHB	0.9635	1.3984	1011.7	−1.402	512
	1.2750	1.4358		4.688	
				1.944	
BPHB	0.9578	2.2117	6284.1	−1.244	491
	1.3858	1.6606		8.350	
				5.306	
MPAB	0.8973	1.0799	151.5	−0.509	157
	0.9780	0.9940		1.709	
				−0.339	
EPAB	0.8932	1.0742	104.7	−0.536	261
	1.1427	1.1458		2.134	
				−0.399	
PPAB	0.9401	1.0559	610.5	−1.317	256
	1.2884	1.3831		2.427	
				−0.496	
BPAB	0.8882	1.2689	1806.3	−2.073	475
	1.3781	1.6007		3.597	
				2.399	

[a]Solutes are: MPHB, methyl *p*-hydroxybenzoate; EPHB, ethyl *p*-hydroxybenzoate; PPHB, propyl *p*-hydroxybenzoate; BPHB, butyl *p*-hydroxybenzoate; MPAB, methyl *p*-aminobenzoate; EPAB, ethyl *p*-aminobenzoate; PPAB, propyl *p*-aminobenzoate; BPAB, butyl *p*-aminobenzoate.[b]Curve-fit parameters are ordered as slope and intercept.[c]Curve-fit parameters are ordered as S_0, S_1 and S_2.

systems exhibit the larger deviations from log-linear behavior, i.e. from $\ln x_A^{sat} = \phi_C^o \ln(x_A^{sat})_C + \phi_W^o \ln(x_A^{sat})_W$. It is not obvious to me how the double log-log will permit accurate mathematical representation of the more non-ideal, non-linear systems when its descriptive ability appears to decrease rapidly with increasing deviation from log-linear behavior.

Mathematical representations should contain provisions for as many curve-fit parameters as is necessary to 'accurately' describe the actual measured data. Moreover, in the case of thermodynamic quantities, mathematical representations should be both continuous functions and differentiable at all points in order to satisfy constraints imposed by thermodynamic relationships such the Gibbs-Duhem equation. Today's computer technology and readily available commercial software packages have made linearization of solubility data an unimportant consideration as far as mathematical representations are concerned, except in those instances when the data set contains only a limited number of experimental values.

References

Acree, W.E., Jr., Mathematical representation of thermodynamic properties. Part 2. Derivation of the combined nearly ideal binary solvent (NIBS)/-Redlich-Kister mathematical representation from a two-body and three-body interactional mixing model. *Thermochim. Acta*, 198 (1992) 71–79.

Acree, W.E., Jr., *IUPAC Solubility Data Series: Polycyclic Aromatic Hydrocarbons in Pure and Binary Solvents*, Oxford University Press, Oxford, UK, 1994, pp. xxxviii–xlii.

Acree, W.E., Jr. and Bertrand, G.L., Thermochemical investigations of nearly ideal binary solvents. 3. Solubility in systems of nonspecific interactions. *J. Phys. Chem.*, 81 (1977) 1170–1173.

Acree, W.E., Jr. and Zvaigzne, A.I., Thermodynamic properties of nonelectrolyte solutions. Part 4. Estimation and mathematical representation of solute activity coefficients and solubilities in binary solvents using the NIBS and modified Wilson equations. *Thermochim. Acta*, 178 (1991) 151–167.

Barzegar-alali, M. and Hanaee, J., Model for solubility estimation in mixed solvent systems. *Int. J. Pharm.*, 109 (1994) 291–295.

Gupta, B., Mishra, D.S., Cheng, C.-H. and Yalkowsky, S.H., Solubility of anthracene in complex solvent systems. *Toxicol. Environ. Chem.*, 33 (1991) 7–21.

Williams, N.A. and Amidon, G.L., Excess free energy approach to the estimation of solubility in mixed solvent systems: I. Theory. *J. Pharm. Sci.*, 73 (1984) 9–13.

Yalkowsky, S.H. and Roseman, T.J., Solubilization of drugs by cosolvents. In Yalkowsky, S.H. (Ed.), *Techniques of Solubilization of Drugs*, Dekker, New York, 1981, pp. 91–134.