ELSEVIER

International Journal of Pharmaceutics 144 (1996) 127–130

international journal of pharmaceutics

Comparison of double log–log, mixture response–surface and combined NIBS/Redlich-Kister solubility models

M. Barzegar-Jalali[a,*], A. Jouyban-Gharamaleki[b], J. Hanaee[b], M.H. Chookhachizadeh-Moghaddam[c]

[a]*Division of Pharmaceutics, School of Pharmacy, Tabriz University of Medical Sciences, Tabriz, Iran*
[b]*Division of Pharmaceutical Chemistry, School of Pharmacy, Tabriz University of Medical Sciences, Tabriz, Iran*
[c]*Department of Biochemistry, University of Payamenur, Ardabil, Iran*

Received 30 June 1996; accepted 28 August 1996

Abstract

The accuracy and prediction capability of the linear double log–log (LDL–L), mixture response–surface (MR–S) and the combined nearly ideal binary solvent/Redlich-Kister (CNIBS/R-K) solubility equations have been compared using the model parameters calculated from either the whole data or a minimum number of data in an experimental set. The CNIBS/R-K model produced better prediction for some experimental sets than the other two models when the parameters obtained from the whole data in a set were employed, whereas the LDL–L model was superior to the other models when the parameters calculated from a minimum number of data were used, indicating its greatest prediction capability.

Keywords: Solubility prediction; Cosolvency; Binary solvent systems; Linear double log–log model; Mixture response–surface method; Combined nearly ideal binary solvent/Redlich-Kister equation

Several models are available for the calculation and prediction of a solute solubility in a binary solvent system, most of which have been cited in a recent report (Barzegar-Jalali and Jouyban-Gharamaleki, 1996).

The purpose of this communication is to compare the accuracy and prediction capability of three of the models, i.e. mixture response–surface (MR–S) (Ochsner et al., 1985), the combined nearly ideal binary solvent/Redlich-Kister (CNIBS/R-K) (Acree and Zvaigzne, 1991) and linear double log–log (LDL–L) (Barzegar-Jalali and Hanaee, 1994) equations using some experi-

* Corresponding author.

0378-5173/96/$15.00 Copyright © 1996 Elsevier Science B.V. All rights reserved
PII S0378-5173(96)04717-5

Table 1
Curve fitting parameters of Eq. (1)

SN[a]	Eq. (1)[b]					Eq. (1)[c]				
	β_1	β_2	β_3	β_4	β_5	β''_1	β''_2	β''_3	β''_4	β''_5
1	−1.60810	−8.80075	0.01171	−0.01261	−0.12200	−0.91229	−8.37991	0.00397	−0.02520	−2.16580
2	−1.77985	−9.88545	0.01143	−0.01235	−0.75420	−0.94392	−9.55603	0.00555	−0.02768	−2.82300
3	−1.27399	−10.75279	0.01200	−0.01938	−2.11370	−1.03910	−10.44904	0.00592	−0.02355	−2.60260
4	0.66216	−11.24471	0.01406	−0.04300	−4.07220	3.03060	−10.30698	−0.00233	−0.08633	−10.45790
5	−3.08436	−9.22003	0.00751	−0.00482	−0.26180	−2.99727	−9.05055	0.00411	−0.00633	−0.48660
6	−2.41775	−9.71974	0.00914	−0.00636	−0.30470	−2.26327	−9.50857	0.00490	−0.00909	−0.62280
7	−1.93956	−10.56784	0.00971	−0.00726	−1.20930	−1.74244	−10.37561	0.00591	−0.01079	−1.60480
8	−1.34613	−11.49043	0.00706	−0.01630	−2.99580	−1.49196	−11.32705	0.00355	−0.01343	−2.72290
9	10.28889	2.39660	0.00950	−0.00756	0.17700	10.72276	2.79781	0.00186	−0.01535	−1.18220
10	9.79422	2.55461	0.00571	−0.00712	2.22860	9.97519	2.82648	0.00043	−0.01026	1.54930
11	12.19089	2.69653	0.00268	−0.01633	0.69420	12.95529	2.99150	−0.00244	−0.03032	−1.39120

[a] SN: System number. SN 1–8 represent methyl *p*-hydroxybenzoate, ethyl *p*-hydroxybenzoate, propyl *p*-hydroxybenzoate, butyl *p*-hydroxybenzoate, methyl *p*-aminobenzoate, ethyl *p*-aminobenzoate, propyl *p*-aminobenzoate and butyl *p*-aminobenzoate in propylene glycol:water mixture, respectively. Data taken from Rubino and Obeng (1991). SN 9–11 denote phenytoin in propylene glycol:water, 1,3-butandiol:water and polyethylene glycol 200:water mixtures. Data taken from Rubino et al. (1984).
[b] The model parameters calculated from whole data.
[c] The model parameters calculated from five data points ($f_c = 0, 0.1, 0.5, 0.7, 1$).

mental data.The MR–S equations used in this report are:

$$\ln X_m = \beta_1 f'_c + \beta_2 f'_w + \left(\frac{\beta_3}{f'_c}\right) + \left(\frac{\beta_4}{f'_w}\right) + \beta_5 f'_c f'_w \quad (1)$$

$$\ln X_m = \beta'_1 f'_c + \beta'_2 f'_w + \beta'_3 f'_c f'_w \quad (2)$$

in which X_m is the solute solubility in the mixed solvent system (cosolvent + water), β_1, β_2, β_3, β_4, β_5, and β'_1, β'_2, β'_3, are the model parameters and f'_c and f'_w are given by $f'_c = 0.96f_c + 0.02$ and $f'_w = 0.96f_w + 0.02$, where f_c and f_w are the volume fractions of the cosolvent and water in the mixture in the absence of the solute. The superiority of Eq. (1) to the extended Hildebrand solubility equation in predicting methylxanthine solubilities in a dioxane–water system has been demonstrated (Ochsner et al., 1985).

The CNIBS/R-K equations employed are:

$$\ln X_m = f_c \ln X_c + f_w \ln X_w + f_c f_w[S_0 + S_1(f_c - f_w) + S_2(f_c - f_w)^2] \quad (3)$$

$$\ln X_m = f_c \ln X_c + f_w \ln X_w + S'_0 f_c f_w \quad (4)$$

X_c and X_w denote the solute solubilities in the neat cosolvent and water, respectively, and S_0, S_1, S_2 and S'_0 are the model parameters. Eq. (3) produced better prediction for some cases than the LDL–L method when the whole data in each experimental set was used for the calculation of the values of S_0, S_1 and S_2 (Acree, 1996).The LDL–L method is expressed by Eqs. (5) and (6):

$$\ln\left[\ln\left(\frac{X_m}{X_w}\right)\right] = \ln\left\{\ln\left[\frac{(X_m)_{0.5}}{X_w}\right]\right\} + B \ln\left(\frac{f_c}{f_w}\right) = \text{Intercept} + \text{Slope} \ln\left(\frac{f_c}{f_w}\right) \quad \text{when } 0 < f_c \leq 0.5 \quad (5)$$

$$\ln\left[\ln\left(\frac{X_c}{X_m}\right)\right] = \ln\left\{\ln\left[\frac{X_c}{(X_m)_{0.5}}\right]\right\} + b \ln\left(\frac{f_w}{0.5}\right) = \text{Intercept} + \text{Slope} \ln\left(\frac{f_w}{0.5}\right) \quad \text{when } 0 < f_w \leq 0.5 \quad (6)$$

where $(X_m)_{0.5}$ is the solute solubility in a system containing 0.5 volume fraction of the cosolvent and/or water. This method was used to linearize the solubility data which had not been linearized

Table 2
Summed squared percentage deviation, $\Sigma(\%D)^2$, for Eqs. (1)–(6)

SN	Whole data points[a]			Five data points[b]			Three data points		
	Eq. (1)	Eq. (3)	Eqs. (5) and (6)	Eq. (1)	Eq. (3)	Eqs. (5) and (6)	Eq. (2)[c]	Eq. (4)[d]	Eq. (5)[c]
1	1393	318	220 (289)[e]	5297	399	143	92	5074	80
2	1673	620	643 (566)	7930	791	721	248	5425	474
3	1751	512	585 (1012)	5001	1968	248	832	13 088	61
4	6091	491	2567 (6284)	148 469	1340	1269	955	13 871	74
5	451	157	113 (152)	1101	443	131	124	2569	114
6	787	261	94 (105)	2142	827	107	232	4461	38
7	734	256	552 (611)	1977	622	599	47	4424	331
8	653	475	998 (1806)	1692	1453	725	519	4697	136
9	1261	664	332 (414)	3636	1469	398	1139	4481	241
10	619	326	260 (373)	1449	888	272	939	2040	257
11	550	18	721 (934)	4486	50	792	218	689	749

[a] $\Sigma(\%D)^2$ values for the whole data were calculated using parameters obtained from data points at f_c values of 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.
[b] $\Sigma(\%D)^2$ values for the whole data were calculated using parameters obtained from data points at f_c values of 0, 0.1, 0.5, 0.7 and 1.
[c] $\Sigma(\%D)^2$ values for cosolvent volume fractions up to 0.5 were calculated using parameters obtained from data points at f_c values of 0, 0.1 and 0.5.
[d] $\Sigma(\%D)^2$ values for cosolvent volume fractions up to 0.5 were calculated using parameters obtained from data points at f_c values of 0, 0.5 and 1.
[e] Values between parentheses are $\Sigma(\%D)^2$ calculated from experimentally determined X_w, $(X_m)_{0.5}$ and X_c.

by the log-linear model and was more accurate than the excess free energy model (Barzegar-Jalali and Hanaee, 1994).

The accuracy and prediction capability of the models have been assessed using Eq. (7):

$$\sum(\%D)^2 = \sum\left\{100\left[\frac{(X_m)_{cal} - (X_m)_{exp}}{(X_m)_{exp}}\right]\right\}^2 \qquad (7)$$

where $\Sigma(\%D)^2$ is the sum of squares of the percent difference between the model predicted and experimentally obtained values of X_m relative to its experimental value at each f_c, $(X_m)_{cal}$ and $(X_m)_{exp}$ denote the predicted and experimental solubilities (X_m) at f_c. It is obvious that the lower $\Sigma(\%D)^2$ is, the more accurate is the model.

The three methods, i.e. Eqs. (1), (3), (5) and (6), were applied to the solubility data of alkyl *p*-hydroxybenzoates and alkyl *p*-aminobenzoates in propylene glycol:water mixtures (Rubino and Obeng, 1991), as well as to phenytoin solubility data in propylene glycol:water, 1,3-butandiol:water, and polyethylene glycol 200:water mixtures (Rubino et al., 1984), using either whole data or a minimum number of data (five data) from each set to obtain the model parameters. The parameters for Eq. (1) obtained from the whole data and five data points are provided in Table 1. For the sake of space, the parameters of the other two models are not given here since most of these have already been reported (Barzegar-Jalali and Hanaee, 1994; Acree, 1996). The parameters obtained in this manner were employed to calculate the values of $(X_m)_{cal}$ for assessing the accuracy of the models according to Eq. (7). The results are seen in Table 2.

It is evident from Table 2 that when the whole data points in a set are used, the accuracy of the models decreases in the following order:

CNIBS/R-K > LDL–L > MR–S

In a previous report (Barzegar-Jalali and Hanaee, 1994), the values of $\Sigma(\%D)^2$ for the benzoates were calculated from the experimentally determined $(X_m)_{0.5}/X_w$ and $X_c/(X_m)_{0.5}$, which are

the values given between parentheses in Table 2, but substitution of the theoretical intercept values for the latter values lowered Σ $(\%D)^2$ and thus enhanced the prediction capability of the LDL–L model.

When five data points were used, the accuracy of the models in decreasing order was as follows:

LDL–L > MR–S > CNIBS/R-K

For five data point combinations other than that given in Table 2, the ranking of the models was generally the same as shown above.

Because of toxicity considerations, the concentration of commonly used cosolvents in the liquid pharmaceutical formulations should be kept as low as possible and should not usually exceed 50% (v/v) (Spiegel and Noseworthy, 1963; Patel et al., 1986; Price et al., 1986; Tsai et al., 1986; Walking et al., 1986; Golightly et al., 1988; Rubino, 1990; USP, 1995). Moreover, in the preformulation stage of a new drug, due to scarcity of the drug, a minimum number of solubility experiments should be conducted in order to predict the solubility at other concentrations of a cosolvent. Therefore, a minimum number of three data points was used to calculate the model parameters so as to verify the prediction capability of the models up to 0.5 volume fraction of cosolvent. From this point of view, Eq. (5) was the best, followed by Eqs. (2) and (4), respectively, indicating its superiority to the other two models. The same was generally true for the three data point combinations other than that given in Table 2.

The analyses given above indicated that no single model was superior in all aspects of accuracy and prediction requirements. For example, the model CNIBS/R-K was the best when a high number of data was employed, whereas the LDL–L model was the most accurate when a minimum number of data was used.

References

Acree Jr., W.E., Comments concerning "Model for solubility estimation in mixed solvent systems". *Int. J. Pharm.*, 127 (1996) 27–30.

Acree Jr., W.E. 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) 151–267.

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

Barzegar-Jalali, M. and Jouyban-Gharamaleki, A., Models for calculating solubility in binary solvent systems. *Int. J. Pharm.*, 140 (1996) 237–246.

Golightly, L.K., Smolinske, S.S, Bennett, M.L., Sutherland III, E.W. and Rumack, B.H., Pharmaceutical excipients adverse effects associated with inactive ingredients in drug products (part I). *Med. Toxicol.*, 3 (1988) 128–165.

Ochsner, A.B., Belloto Jr., R.J. and Sokoloski, T.D., Prediction of xanthine solubilities using statistical techniques. *J. Pharm. Sci.*, 74 (1985) 132–135.

Patel, D.M., Bernardo, P., Cooper, J. and Forrester, R.B., Glycerin. In *Handbook of Pharmaceutical Excipients*, American and Great British Pharmaceutical Societies, London, 1986, pp. 123–124.

Price, J.C., Gore, A.Y., Chowhan, Z. and Horder, R.L., Polyethylene glycol. In *Handbook of Pharmaceutical Excipients*, American and Great British Pharmaceutical Societies, London, 1986, pp. 209–213.

Rubino, J.T., Cosolvents and cosolvency. In Swarbrick, J. and Boylan, J.C. (Eds), *Encyclopedia of Pharmaceutical Technology*, Marcel Dekker, New York, 1990, pp. 375–398.

Rubino, J.T., Blanchard, J. and Yalkowsky, S.H., Solubilization by cosolvents: II. Phenytoin in binary and ternary solvents. *J. Parenter. Sci. Technol.*, 38 (1984) 215–221.

Rubino, J.T. and Obeng, E.K., Influence of solute structure on deviation from log-linear solubility equation in propylene glycol:water mixtures. *J. Pharm. Sci.*, 80 (1991) 479–483.

Spiegel, A.J. and Noseworthy, M.M., Use of nonaqueous solvents in parenteral products. *J. Pharm. Sci.*, 52 (1963) 917–927.

Tsai, P.S., Lipper, R.A. and Worthington, H.C., Propylene glycol. In *Handbook of Pharmaceutical Excipients*, American and Great British Pharmaceutical Societies, London, 1986, pp. 241–242.

USP, 23rd Revision, US Pharmacopeial Convention, Rockville, MD, 1995, pp. 220–221, 469, 480, 515, 659, 683, 1189, 1202, 1261, 1369, 1402, 1528, 1604, 2215, 2218.

Walking, D., Rowe, E.L. and Worthington, H.C., Glycofural. In *Handbook of Pharmaceutical Excipients*, American and Great British Pharmaceutical Societies, London, 1986, p. 127.