Review of the cosolvency models for predicting solubility of drugs in water-cosolvent mixtures

Abolghasem Jouyban

Faculty of Pharmacy and Drug Applied Research Center, Tabriz University of Medical Sciences, Tabriz, Iran.

ABSTRACT - The cosolvency models presented from 1960 to 2007 were reviewed and their accuracies for correlating and/or predicting the solubility of drugs in water-cosolvent mixtures were discussed. The cosolvency models could be divided into theoretical, semi-empirical and empirical models, the first group of models provide basic information from the solution, while the last group of models are suitable for solubility correlation studies. The simplest cosolvency model, i.e. the log-linear model of Yalkowsky, provides an estimate of drug solubility in water-cosolvent mixtures using aqueous solubility of the drug, whereas the Jouyban-Acree model predicts the solubility with an acceptable error with the cost of one more data point (the solubility in neat cosolvent) which is required as input value in the prediction process. A number of error terms used in the literature was also discussed with a brief comments on the acceptable prediction error for pharmaceutical applications.

INTRODUCTION

The knowledge of solubility is important in the pharmaceutical area, because it permits the scientist the choice of the best solvent medium for a drug or combination of drugs, and helps in overcoming certain difficulties arising in the preparation of pharmaceutical solutions. These solutions could be used to test the purity of bulk drugs, to prepare a liquid dosage form and/or to extract an ingredient from a synthetic mixture or a natural source. A detailed investigation of the solubility phenomenon and related properties also yields useful information about the structure and intermolecular forces of drugs. The solubility of a drug depends on the physical and chemical properties of the drug and the solvent including the polarity, dielectric constant, autoprotolysis constant of the solvent (pKs) and also factors such as temperature and pH of the solution. Water is the main solvent and the aqueous solubility is one of the most important properties of a drug molecule. The drugs are often low soluble in water and their solubility should be increased. A comprehensive database of aqueous solubility data of chemicals and pharmaceuticals was collected by Yalkowsky and He (1). There are several methods to enhance the aqueous solubility of the drugs including cosolvency, hydrotropism, complexation, ionisation and using the surface active agents. These methods were discussed in details (2).

The cosolvency, mixing a permissible non-toxic organic solvent with water, is the most common technique to increase the aqueous solubility of drugs. The common cosolvents which, are used in the pharmaceutical industry are ethanol, propylene glycol, glycerine, glycofural and polyethylene glycols (3). Their applications and possible side effects were discussed in the literature (as examples see 4-8). The cosolvency phenomenon has wide ranging applications in different fields. In addition to applications in drug formulation, solubilizing agents like tert-butyl ether were used clinically to dissolve cholesterol gallstones (9). From the environmental scientist's viewpoint, the cosolvency is an important subject, because the organic solvents can change distribution and movement of hydrophobic contaminants in the environment (10). The solubility behaviour of a solute in mixed solvents provides some theoretical basis for the chemist. From this, it is possible to collect some ideas about solute-solvent and solvent-solvent interactions in the solution.

Cosolvency of non-aqueous solvent mixtures is also important from a pharmaceutcial viewpoint, since the mixtures could be used as synthesis medium or re-crystalisation solvents for purification of drugs.

Correspondence Author: Prof. A. Jouyban, Faculty of Pharmacy, Tabriz University of Medical Sciences, Tabriz 51664, Iran, Email: ajouyban@hotmail.com

The addition of a cosolvent to the aqueous solution can also affect the chemical stability of the drug (11), acid dissociation constants of the solute, viscosity and surface tension of the solution which should be kept in mind in pharmaceutical applications of the cosolvents. The aims of this work is to review the available cosolvency models, their prediction capabilities and also the error ranges of the predicted solubilities in the pharmaceutical areas. The basic models and their new extensions were discussed along with the available accuracy test of the models.

REVIEW OF COSOLVENCY MODELS

Apart from experimental determinations of solute solubility in water-cosolvent mixtures, there are many mathematical models describing the solute solubility in mixed solvents (12-23). By considering chemical theory, some of them are theoretical and some others are semi-theoretical or empirical. Theoretical models provide some evidence for better understanding of solubility behaviour for drugs in mixed solvents, while semi-theoretical or empirical approaches are very useful models for correlating experimental solubilities to the independent variables such as volume fraction of the cosolvent.

From a practical point of view, one can categorise these models into two groups, i.e. predictive and correlative models. The predictive cosolvency model means that the model is capable of predicting the solubility of a drug in mixed solvents by using no experimental solubility data (pure or fully predictive model) and/or using one or two solubility points as input data. The correlative cosolvency model means that the model provides good correlation between the solubility of a drug in mixed solvents and the independent variables like cosolvent concentration or physico-chemical properties of the solutes or cosolvents. The advantage of the predictive models such as universal functional group activity coefficient, UNIFAC, (24) is that these models do not employ any experimental data points. However, the low prediction capability of the UNIFAC model for biphenyl solubilities in several binary mixed solvents (as a simple model system) has been reported (25). In addition, these methods usually require a number of physico-chemical properties of drugs and relatively complex computations and are not more favourable methods in pharmaceutical area. The correlative equations, such as the general

single model (22), employ the curve-fitting parameters to correlate experimental solubility data with respect to the concentration of the cosolvent. In order to calculate these curve-fitting parameters, one has to determine a set of experiments in mixed solvents to train the model. From a practical point of view, a model containing a minimum number of the curve-fitting parameters is the best cosolvency model.

The final goal of developing cosolvency equations is that it enables researchers to predict the solute solubility in mixed solvents from a minimum number of experiments or even without experimental data. It has been shown that solubility prediction in binary solvents using correlative equations can fail when using insufficient experimental data points and produces unacceptable errors (22). On the other hand, Bustamante and co-workers (26) employed a modified form of the extended Hildebrand solubility approach to correlate structurally related drugs solubility in binary solvent mixtures. The authors used the solute solubility in water and cosolvent, the solute's solubility parameter, the Hildebrand solubility parameter of the solvent and the basic solubility parameter of the mixed solvent as independent variables. This approach was a useful solution to the solubility problem of similar chemically solutes. The applicability of the Jouyban-Acree model for reproducing solubility data of structurally related drugs in binary solvents was presented (27). It was also shown that the prediction error of the Jouyban-Acree model is less than that of the modified form of the extended Hildebrand solubility approach (27).

HILDEBRAND SOLUBILITY APPROACH

A solution obeying Raoult's law is known as an ideal solution and its solubility could be calculated from the heat of fusion of the solute (ΔH_m^f) and the differenec in the heat capacities of this solid and of its supercooled liquid (ΔC_p) using Hildebrand and Scott equation expressed by:

$$\log X = -\frac{\Delta H_m^f}{4.575}\left(\frac{T_m - T}{T \cdot T_m}\right) + \frac{\Delta C_p}{4.575}\left(\frac{T_m - T}{T}\right) - \frac{\Delta C_p}{1.987}\log\left(\frac{T_m}{T}\right)$$

(1)

where X is the mole fraction solubility at temperature T, T_m is the melting point of the solute,

$\Delta C_p = C_p^l - C_p^s$ where C_p^l and C_p^s are the molal heat capacities of the liquid and solid forms, respectively. For regular solutions, the equation was modified as:

$$\log X = -\frac{\Delta H_m^f}{4.575}\left(\frac{T_m - T}{T \cdot T_m}\right) + \frac{\Delta C_p}{4.575}\left(\frac{T_m - T}{T}\right)$$
$$-\frac{\Delta C_p}{1.987}\log\left(\frac{T_m}{T}\right) - \frac{V_s \varphi_m}{4.575T}\left(\delta_m - \delta_s\right)^2$$
$$(2)$$

in which V_s is the molar volume of the solute, ϕ_m is the volume fraction of the solvent, δ_m and δ_s are the solubility parameters of the solvent and solute, respectively. Although Hildebrand restricts the application of the model to nonpolar solvents, Chertkoff and Martin (12) used the model for calculating the solubility of benzoic acid in binary mixtures of hexane, ethyl acetate, ethanol and water. These mixtures provided wide polarity range of solvent from 7.3 H (for hexane) to 23.4 H (for water) and the maximum solubility of benzoic acid was observed in δ_m =11.5 H (H: Hildebrand unit, 1

$$H= 1\left(\frac{cal}{cm^3}\right)^{\frac{1}{2}} = 0.489 \text{ MPa}^{1/2}).$$

The numerical values of δ_m of binary solvents were calculated using $\delta_m = f_1\delta_1 + f_2\delta_2$ in which f_1 and f_2 are the fractions of solvents 1 (cosolvent) and 2 (water), δ_1 and δ_2 are the solubility parameters of the solvents 1 and 2. This equation could be rearranged as a linear relationship by replacing f_2 with $1 - f_1$ as $\delta_1 = \delta_2 + (\delta_1 - \delta_2)f_1 = Intercept + slope \cdot f_1$. On the other hand, all terms of equation (2) are constant values for a given solute at a fixed temperature except than δ_m, therefore, the equation could be converted as a simple linear relationship of $\log X = Intercept + slope \cdot f_1$ (22). This model was used by Yalkowsky *et al.* and was called log-linear model.

SOLUBILITY – DIELECTRIC CONSTANT RELATIONSHIP MODEL

Paruta and co-workers tried to correlate the solubility of salicylic acid as a function of dielectric constant (ε) of the solvent mixture (13) where ε values were determined using a resonance method. As shown in Figure 1, the model predict the maximum solubility of the solute in binary solvents, however, there are different observed solubilities for a given ε value. This means that the dielectric constant can not adequately represent the solvent effects on the solubility of solutes.

This apporach was employed later on as a polynomial of ε values to correlate the logarithm of solubilities in mixed solvents (28). The numerical values of the ε_m were calculated using $f_1\varepsilon_1 + f_2\varepsilon_2$ (29) which is not correct calculations, since there is a non-linear relationship between experimental ε_m and f_1 values as was shown in a paper (30). However, the main reason for the accurate correlation of the $\ln X_m$ using ε_m polynomial is that the polynomial could be arranged as a polynomial of f_1 values by simple algebraic manipulations (31).
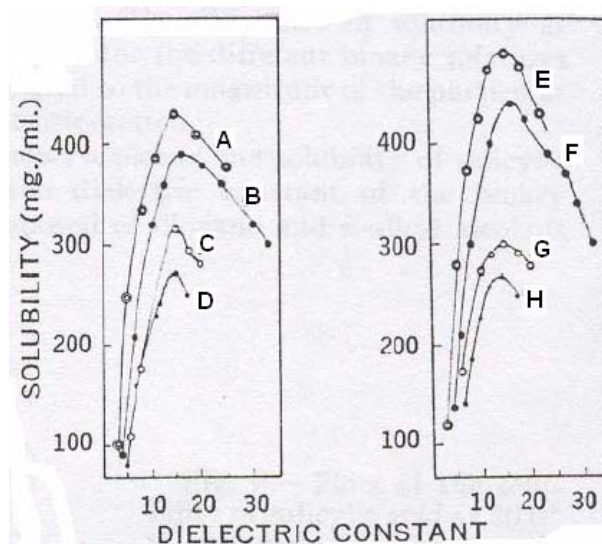

Figure 1. Plot of solubility of salicylic acid at 30.6 °C in binary mixtures of benzene with ethanol (A), methanol (B), 1-propanol (C), 1-butanol (D) and binary mixtures of dioxane with ethanol (E), methanol (F), 1-propanol (G), 1-butanol (H), (Reproduced from a reference (13)).

THE LOG-LINEAR MODEL OF YALKOWSKY

The algebraic mixing rule (15) or log-linear model which is expressed by:

$$\log X_m = f_1 \log X_1 + f_2 \log X_2 \qquad (3)$$

where X_m is the mole fraction solubility of the solute, X_1 and X_2 denote the mole fraction solubility in neat cosolvent and water, respectively. The model could be rearranged as:

$$\log X_m = \log X_2 + \left(\log \frac{X_1}{X_2} \right) f_1$$
$$= Interept + Slope \cdot f_1 \qquad (4)$$

Equation (4) is a correlative model, however, it has been demonstrated that a log-linear relationship between the solubility of a non-polar solute and the fraction of the cosolvent exist as (15):

$$\log X_m = \log X_2 + \sigma \cdot f_1 \qquad (5)$$

where σ is the solubilization power of the cosolvent and theoretically is equal to $\left(\log \frac{X_1}{X_2} \right)$.

Valvani *et al.* (32) reported a linear relationship between σ and logarithm of drug's partition coefficient ($\log K_{ow}$) which is a key relationship and could improve the prediction capability of the log-linear model. The relationship was expressed as:

$$\sigma = M \cdot \log K_{ow} + N \qquad (6)$$

where M and N are the cosolvent constants and are not dependent on the solute's nature. The numerical values of M and N were reported for most of the common cosolvents earlier (33) and summarized in Table 1. This version of the log-linear model could be considered as a predictive model and provided the simplest solubility estimation method and requires the aqueous solubility of the drug and its experimental/calculated logP value as input data. Updated M and N values for ethanol (0.93, 0.40), propylene glycol (0.77, 0.58), polyethylene glycol

400 (0.74, 1.26) and glycerol (0.35, 0.26) were reported by Millard *et al.* (34) employing published data and the data selection criteria of: a) adequate time for equilibration (24 h) or test for equilibration, b) room temperature, i.e. 22-27 °C, experiments and c) at least duplicated experiments. These slight variations of M and N values could not affect the prediction capability of the log-linear model as shown in a recent work (35).

With the known values of M and N along with the $\log K_{ow}$ of a drug, it is possible to predict the cosolvent concentration for solubilization of the desired amount of the drug employing only the experimental aqueous solubility data (34). This prediction method produces average absolute error (*AAE*) of ~ 0.5 (in log scale) for the solubility of drugs in water-ethanol mixtures and could be used employing experimental values of $\log K_{ow}$ or computed values using various software (35). The *AAE* of ~0.5 (for 26 data sets reported in the reference (35) is equal to the *MPD* of ~ 300 % (Unpublished results). The *MPD* values are defined similar to the relative standard deviation (RSD) and could be compared with the RSD of the solubilities from the repeated experiments. The produced *MPDs* is relatively high when compared with the experimentally determined RSD values (usually < 10%) and acceptable percentage error range (~ 30 %) in pharmaceutical area (36-37), however, the method is strightforward and easy to use and employs only aqueous solubility data of the drug (one data point).

Li (38) extended the log-linear model using the activity coefficients of water and cosolvent computed by UNIFAC method. The most accurate extended model of Li was:

$$\log X_m = \log X_2 + \sigma \cdot f_1 + \ln \alpha_2 + \phi_1 \ln \left(\frac{\alpha_1}{\alpha_2} \right)$$
$$(7)$$

Where α_2 and α_1 are the activity coefficients of solvents 2 (water) and 1 (cosolvent) in the mixture, ϕ_1 is the mole fraction of the cosolvent in the binary solvent mixture in the absence of the solute. Li compared the accuracy of the extended model with the log-linear model using average error of prediction (*AEP*) calculated by:

$$AEP = \frac{\sum \left| \log X_m^{Calculated} - \log X_m^{Observed} \right|}{N}$$
(8)

Table 2 listed the AEP of the log-linear and the extended model for the predicted solubilities of various solutes dissolved in 13 different water-cosolvent mixtures. This extension was improved the prediction capability of the log-linear model for 7 of the 13 cosolvents (38).

The solubilities in mono-solvents, i.e. $\log X_1$ and $\log X_2$ (aqueous solubility), could be calculated using:

$$\log X_i = \frac{-\Delta S_f (mp - 25)}{1364} - \log P_{\frac{Oc\tan ol}{solvent\ i}} + C_i$$
(9)

where ΔS_f is the entropy of fusion, mp is the melting point of the drug, $P_{\frac{Oc\tan ol}{solvent\ i}}$ is the octanol-solvent i prtition coefficient of the solute and C_i is a constant which is depend upon the unit of the solubility (39). The accuracy of the full predictive version of the log-linear model where $\log X_2$ calculated using equation (9), has not been reported in the literature.

The solubility profile of drugs are often linear up to $f_1 = 0.5$ and it has been shown that the log-linear model could be modified as:

$$\log X_m = \log X_2 + \sigma_{0.5} \cdot f_1$$
(10)

which provides more accurate predictions and is more practical in pharmaceutical area since most of the cosolvents employed in $f_1 < 0.5$ (8) concentrations. There are also linear relationship between $\sigma_{0.5}$ and $\log K_{ow}$ of the solutes and summary of the M and N values were listed in Table 1.

The main assumptions on which the log-linear model is based are as follows: a) the free energy of transferring a solute to an ideal solvent mixture is the sum of the corresponding energies in pure solvents, b) the solvent molecules behave in a mixture the same way as they do in neat solvents, c) the ratio of the solvent and cosolvent surrounding a solute molecule is the same as volume fraction of the solvents in the mixture, d) the molar volume of solute in the solution is not so different from the molar volumes of the solvent and cosolvent, e) no degradation, solvation or solvent mediated polymorphic transitions of the solute occur (40). However, since most of these assumptions are not applicable to the solubility of drugs in aqueous binary mixtures, the model produces relatively large deviations from the true experimental data (see reference 22).

The model was extended to the ternary solvents (41) as:

$$\log X_m = f_1 \log X_1 + f_2 \log X_2 + f_3 \log X_3 = B_0 + B_1 f_2 + B_3 f_3$$
(11)

and for quinary mixtures (42-43) as:

$$\log X_m = f_1 \log X_1 + f_2 \log X_2 + f_3 \log X_3 + f_4 \log X_4$$
$$= B_0 + B_1 f_2 + B_2 f_3 + B_3 f_4$$
(12)

where X_3 and X_4 are the solubilities in neat solvents 3 and 4 and the f terms denote the fractions of the solvents in the mixture. The general form of the log-linear model for multicomponent solvent systems could be written as:

$$\log X_m = \log X_2 + \sum \sigma_i f_i$$
(13)

where σ_i and f_i are the solubilization power and the fractions of cosolvent i (38).

The accuracy of the log-linear model and its extended version for calculating the solubilities in ternary and quinary solvents were tested using three solubility data sets for each solvent systems. The overall *AEP* of the log-linear and the extended model for ternary solvents were 0.247 and 0.203, respectively and the corresponding values for quinary solvents were 0.330 and 0.268 (38).

EXTENDED HILDEBRAND SOLUBILITY APPROACH

By the extended Hildebrand solubility approach (EHS), Martin and coworkers (14) extended the applicability of the regular solution theory to the solubility of drugs in water-cosolvent mixtures by avoiding Hildebrand's geometric assumption for the interaction term.

Table 1. Numerical values of M and N of common cosolvents for calculating the slope (σ) and half-slope ($\sigma_{0.5}$) of the log-linear model (33).

	σ		$\sigma_{0.5}$	
	M	N	M	N
Acetone	1.14	-0.10	1.25	0.21
Acetonitrile	1.16	-0.49	1.04	0.44
Butylamine	0.64	1.86	0.67	3.83
Dimethylacetamide	0.96	0.75	0.89	1.28
Dimethylformamide	0.83	0.92	0.65	1.70
Dimethylsulphoxide	0.79	0.95	0.72	0.78
Dioxane	1.08	0.40	0.99	1.54
Ethanol	0.95	0.30	0.81	1.14
Ethylene glycol	0.68	0.37	0.52	0.28
Glycerol	0.35	0.28	0.38	0.14
Methanol	0.89	0.36	0.73	0.70
Polyethylene glycol 400	0.88	0.68	0.78	1.27
1-Propanol	1.09	0.01	1.03	1.76
2-Propanol	1.11	-0.50	0.96	1.00
Propylene glycol	0.78	0.37	0.55	0.87

Table 2. The average error of prediction (AEP) and standard deviation (SD) of various models for calculating the solutes solubility in the common cosolvents and their overall values taken from a reference (38)

		Log-linear		Extended log-linear	
Cosolvent	N^a	AEP	SD	AEP	SD
Acetone	220	0.306	0.303	0.246	0.310
Acetonitrile	103	0.416	0.284	0.208	0.256
Dimethyl sulfoxide	131	0.186	0.253	0.316	0.311
Dimethylacetamide	97	0.263	0.312	0.412	0.357
Dimethylformamide	133	0.216	0.260	0.221	0.262
Dioxane	349	0.608	0.334	0.386	0.295
Ethanol	1631	0.328	0.379	0.261	0.357
Ethylene glycol	111	0.186	0.148	0.265	0.164
Glycerol	124	0.096	0.119	0.095	0.117
Methanol	688	0.274	0.375	0.280	0.371
1-Propanol	116	0.570	0.397	0.356	0.336
2-Propanol	188	0.413	0.413	0.291	0.366
Propylene glycol	503	0.193	0.295	0.212	0.292
Overall		0.318	0.392	0.269	0.362

^{a} N is the number of data points.

In the original Hildebrand equation the solute-solvent interaction term is assumed equal to ($\delta_m \times \delta_s$) in which δ_m and δ_s are the solubility parameters of mixed solvent and solute, respectively, and the model can describe the regular behaviour of the solution. Instead Martin's group used an empirical solute-solvent interaction parameter (*WW*). This modification widened the applications of the model to semi-polar crystalline drugs in irregular solutions involving self-association and hydrogen-bonding, such as occurs in polar binary mixtures. Using the EHS model, the co-logarithm of the mole fraction solubility ($-\log X_m$) can be expressed as:

$$-\log X_m = -\log X_s^i + \frac{V_s \phi_m^2 \left(\delta_m^2 + \delta_s^2 - 2WW \right)}{2.303RT}$$
(14)

where X_s^i denotes the ideal mole fraction solubility of the solute, V_s is the molar volume of the solute, ϕ_m represents the volume fraction of the solvent in solution and because of very low solubility of the solute it can be assumed equal to 1 (29, 44-46), R is the molar gas constant, T denotes absolute temperature and *WW* is the interaction term whichis calculated by a power series of δ_m:

$$WW = \sum_{i=0}^{p} A_i \delta_m^i$$
(15)

where A_i denotes the curve-fit parameter and δ_m is calculated by using equation (16):

$$\delta_m = f_1\delta_1 + f_2\delta_2$$
(16)

in which δ_1 and δ_2 are the solubility parameters of pure cosolvent and water, respectively (26). However, in order to obtain an estimation of ideal solubility based on experimentally determined entropy or enthalpy of fusion, a sophisticated high cost instrument, such as a differential scanning calorimeter is required for the measurements. In addition to the experimentally measured ideal solubility, solution density and an estimation of physical parameters of V_s and δ_s are required. All are essential to calculate solute solubility by the EHS equation. In addition, the dependence of δ_s values on solvent polarity restricts the applications of EHS for predictive purposes (27).

The model was modified to directly relate the solubility to the solubility parameters of the solvent mixtures (26, 36, 47) as:

$$\log X_m = C_0 + C_1\delta_m + C_2\delta_m^2 + C_3\delta_m^3 + \cdots + C_n\delta_m^n$$
(17)

where C terms are the curve fitting parameters. Using this modified version there are no need for experimental determination of ideal solubility of the solute and other terms required in the extended Hildebrand solubility apaproach (26). This polynomial, i.e. equation (17), was coverted to the GSM using simple algebraic manipulations (22). This version of the EHS is a correlative model and did not provide accurate predictions.

An extended version of Martin's EHS model was proposed for describing the multiple solubility maxima of solutes in solvent mixtures where the WW term was correlated with a power series of the solvent compositions. This extension was applied to describe the multiple solubility maxima of five drugs in water-ethanol and ethanol-ethyl acetate mixtures with the mean percentage deviation (*MPD*) of ~ 11 % (48).

THE WILLIAMS-AMIDON MODEL

The excess free energy models of Williams-Amidon (16) are expressed by:

$$\log X_m = f_1 \log X_1 + f_2 \log X_2 + A_{1-2} f_1 f_2 \left(\frac{V_s}{V_1} \right)$$
(18)

$$\log X_m = f_1 \log X_1 + f_2 \log X_2 - A_{1-2} f_1 f_2 (2f_1 - 1) \left(\frac{V_s}{V_1} \right)$$
$$+ 2A_{2-1} f_1^2 f_2 \left(\frac{V_s}{V_2} \right) + C_s f_1 f_2$$
(19)

$$\log X_m = f_1 \log X_1 + f_2 \log X_2 - A_{1-2} f_1 f_2 (2f_1 - 1) \left(\frac{V_s}{V_1} \right)$$
$$+ 2A_{2-1} f_1^2 f_2 \left(\frac{V_s}{V_2} \right) + 3D_{12} f_1^2 f_2^2 \left(\frac{V_s}{V_2} \right) + C_2 f_1 f_2^2 \left(\frac{V_s}{V_2} \right) + C_1 f_1^2 f_2$$
(20)

where A_{1-2}, A_{2-1}, C_s, D_{12}, C_2 and C_1 are solvent-solvent or solute-solvent interaction terms, V_1 and V_2 represent the molar volumes of cosolvent and water, respectively (16). The molar volume differences between solvent, cosolvent and solute are considered as the V terms. The solvent-cosolvent interaction constants (D_{12} and A terms) are obtained from vapour-liquid equilibrium data. Williams and Amidon (16) reported different values

of A for two-suffix (equation (18)), three-suffix (equation (19)) and four-suffix (equation (20)) excess free energy models with respect to a given binary solvent system. This fact is due to different definitions of the A terms in the three equations. The solute-solvent interaction terms are estimated from experimental solubility data. This model is able to improve the predictability of the log-linear model by employing additional terms. The excess free energy models were also extended to express the solubility of drugs in ternary solvent mixtures (49). These models could be considered as predictive models and require solubilities in neat solvents and one datum in mixed solvents as input data to provide predictions.

Equations (18)-(20) can however be simplified to the following forms (50):

$$\log X_m = f_1 \log X_1 + f_2 \log X_2 + \alpha_1 f_1 f_2$$
(21)

$$\log X_m = f_1 \log X_1 + f_2 \log X_2 + \alpha'_1 f_1 f_2 + \alpha'_2 f_1^2 f_2$$
(22)

$$\log X_m = f_1 \log X_1 + f_2 \log X_2 + \alpha''_1 f_1 f_2 + \alpha''_2 f_1^2 f_2 + \alpha''_3 f_1 f_2^2 + \alpha''_4 f_1^2 f_2^2$$
(23)

where $\alpha_1 = [A_{1-2}(V_s/V_1)]$, $\alpha'_1 = [A_{1-2}(V_s/V_1) + C_2]$, $\alpha'_2 = 2[A_{2-1}(V_s/V_2) - A_{1-2}(V_s/V_1)]$, $\alpha''_1 = [A_{1-2}(V_s/V_1)]$, $\alpha''_2 = 2[A_{2-1}(V_s/V_2) - A_{1-2}(V_s/V_1)] + C_1(V_s/V_1)$, $\alpha''_3 = C_3(V_s/V_2)$ and $\alpha''_4 = [3D_{12}(V_s/V_2)]$.

Equation (22) was extended for mathematical representation of solubility data of drugs in binary solvent mixtures at various temperatures using:

$$\log X_m = f_1 \log X_{1,T} + f_2 \log X_{2,T} + W_1 \left(\frac{f_1 f_2}{T} \right) + W_2 \left(\frac{f_1^2 f_2}{T} \right)$$
(24)

where W_1 and W_2 are the model constants representing the solute and solvents 1 and 2 intereactions. The model was trained using a minimum number of five experimental solubility data points of oxolinic acid at 20 and 40 °C, then the solubility of the solute in binary mixtures of ethanol with water and ethyl acetate at 20-40 °C were predicted. The overall average relative error (*ARE*) was ~ 8 % (51). The accuracy of the extended excess free energy model was also

checked using 7 solubility data sets at various temperatures collected from the literature and the overall *ARE* was ~13 %. The corresponding *ARE* for similar computations employing the Jouyban-Acree model was ~10 % (51).

MIXTURE RESPONSE SURFACE MODEL

Statistically based mixture response surface methods, MRS, (17) have been proposed for correlative purposes and these models are as follows:

$$\log X_m = \beta_1 f_1' + \beta_2 f_2' + \beta_3 f_1' f_2' \qquad (25)$$

$$\log X_m = \beta_1' f_1' + \beta_2' f_2' + \beta_3'\left(\frac{1}{f_1'}\right) + \beta_4'\left(\frac{1}{f_2'}\right) \quad (25)$$

\

$$\log X_m = \beta_1'' f_1' + \beta_2'' f_2' + \beta_3''\left(\frac{1}{f_1'}\right) + \beta_4''\left(\frac{1}{f_2'}\right) + \beta_5'' f_1' f_2'$$
$$(26)$$

in which β_1 - β_3, β_1' - β_4' and β_1'' - β_5'' are the model's parameters and f_1' and f_2', are given by f_1' = 0.96 f_1 + 0.02 and f_2' = 0.96 f_2 + 0.02 (17). By converting f values to f', the model can cover the whole range of volume fraction of the cosolvent (f_1: 0-1). The authors showed the superiority of the MRS equations to the EHS model by examining xanthine derivative solubilities in dioxane-water mixture (17). The MRSs are correlative models and no report was published on their prediction capabilities.

THE KHOSSRAVI-CONNORS MODEL

The Khossravi-Connors model (20) was formulated as the summary of the free energy changes of three steps involved in the dissolution of a solute in a solvent system:

$$\Delta G_{total}^0 = \Delta G_{crystal}^0 + \Delta G_{cavity}^0 + \Delta G_{solvation}^0 \qquad (27)$$

where ΔG_{total}^0 is the total free energy change, $\Delta G_{crystal}^0$ is the crystal lattice energy and any solute-solute interactions in the solution representing the conversion of the crystalline solute to the gaseous solute, ΔG_{cavity}^0 is the free energy

change of cavity formation and $\Delta G_{solvation}^0$ is the free energy changes of insertion of the gaseous solute in the cavity and its solvation processes. Khossravi and Connors called $\Delta G_{crystal}^0$ as intersolute effect, ΔG_{cavity}^0 as medium effect and $\Delta G_{solvation}^0$ as the solvation effect. The total free energy change could be related to the experimental solubility of a solute in binary solvent mixtures (X_m) using:

$$\Delta G_{total}^0 = -kT \ln X_m \qquad (28)$$

where k is the Boltzmann's constant (20).
The crystal composition of a solute is independent of the solvent composition, therefore, $\Delta G_{crystal}^0$ (intersolute effect) is a solvent independent term. The authors correlated the ΔG_{cavity}^0 and $\Delta G_{solvation}^0$ to the solvent composition of the aqueous-organic solvent mixtures. The ΔG_{cavity}^0 was formulated as:

$$\Delta G_{cavity}^0 = g \cdot A \cdot \left[\gamma_1 + \gamma'\left(\frac{\beta_1 f_1 f_2 + 2\beta_2 f_2^2}{f_1^2 + \beta_1 f_1 f_2 + \beta_2 f_2^2} \right) \right]$$
$$(29)$$

where g is the curvature effect factor, A is the area defined by the van der Waals radii of the solute atoms and assumed as solvent independent quantity, γ_1 is the surface tension of pure solvent 1, $\gamma' = \frac{\gamma_2 - \gamma_1}{2}$ in which γ_2 is the surface tension of the solvent 2, β_1 and β_2 are functions of the equilibrium constants (K_1 and K_2) of the solvation of the solute in the binary solvent mixture, i.e. $\beta_1 = K_1$ and $\beta_2 = K_1 K_2$. The solvation effect was presented as:

$$\Delta G_{solvation}^0 = \Delta G_{11}^0 F_{11} + \Delta G_{12}^0 F_{12} + \Delta G_{22}^0 F_{22} \quad (30)$$

where F_{11}, F_{12} and F_{22} are fractions of the solute in the solvated forms by two molecules of solvent 1, one molecule of solvent 1 and one molecule of solvent 2 and two molecules of solvent 2, respectively, ΔG_{11}^0, ΔG_{12}^0 and ΔG_{22}^0 are the free energy changes of the solvation processes by 1-1, 1-

2 and 2-2 molecules of the solvents. Equation (30) could be re-written as:

$$\Delta G^0_{solvation} = S_1 F_{12} + S_2 F_{22} + \Delta G^0_{11} \qquad (31)$$

where $S_1 = \Delta G^0_{12} - \Delta G^0_{11}$ and $S_2 = \Delta G^0_{22} - \Delta G^0_{11}$. By considering the equilibrium constants of the exchange equilibria of water by the cosolvent and further simplifications, equation (32) could be obtained:

$$\Delta G^0_{solvation} = \frac{S_1 \beta_1 f_1 f_2 + S_2 \beta_2 f_2^2}{f_1^2 + \beta_1 f_1 f_2 + \beta_2 f_2^2} + \Delta G^0_{11}. \qquad (32)$$

Khossravi and Connors used the Leffler-Grunwald delta operator symbolism and defined the total solvent effects ($\Delta G^0_{total}(f_2)$) as:

$$\delta_M \Delta G^0 = \Delta G^0_{total}(f_2) - \Delta G^0_{total}(f_2 = 0) \qquad (33)$$

and all solvent composition independent quantities were vanished:

$$\delta_M \Delta G^0 = \frac{(-kT\ln\beta_1 + g \cdot A \cdot \gamma)\beta_1 f_1 f_2 + (-kT\ln\beta_2 + 2g \cdot A \cdot \gamma)\beta_2 f_2^2}{f_1^2 + \beta_1 f_1 f_2 + \beta_2 f_2^2}$$
$$(34)$$

The authors evaluated their proposed model employing 44 solubility data sets in 8 different cosolvents at various temperatures. The calculated curve-fit criterion (CFC) was:

$$CFC = 100\left(\frac{SD\ of\ the\ points\ about\ the\ line}{mean\ of\ the\ ordinate\ values}\right)$$
$$(35)$$

and was in the approximate range of 0.8-7.5 %, with the overall of 2.6 % (20). These percentage were computed considering the $\delta_M \Delta G^0$ and the authors did not convert the $\delta_M \Delta G^0$ terms to the molar or mole fraction solubilities to be compared with the corresponding solubilities calculated by other cosolvency models. The model was derived based on thermodynamic approached, however, no report is avaliable on its prediction capability from the literature.

THE JOUYBAN-ACREE MODEL

The Jouyban-Acree model, formerly known as the combined nearly ideal binary solvent/Redlich-Kister equation, was derived from a thermodynamic mixing model that includes contributions from both two-body and three-body interactions. The model was presented for solubility calculations in mixed solvents by our group (19, 52) and was expressed as:

$$\log X_m = f_1 \log X_1 + f_2 \log X_2 + f_1 f_2 \sum_{i=0}^{2} S_i (f_1 - f_2)^i$$
$$(36)$$

where S_i stands for the model constants. The S_i values can be calculated by two procedures:

1)   regressing
$$\left(\frac{\log X_m - f_1 \log X_1 + f_2 \log X_2}{f_1 f_2}\right)$$

against $(f_1 - f_2)$ and $(f_1 - f_2)^2$ by a classical least square analysis (18) and

2)   regressing
$(\log X_m - f_1 \log X_1 + f_2 \log X_2)$ against $f_1 f_2$, $f_1 f_2 (f_1 - f_2)$ and $f_1 f_2 (f_1 - f_2)^2$ by a no intercept least squares analysis. This procedure produced more accurate calculations than the one above for the solute's solubility in aqueous binary solvents (53). The Jouyban-Acree model is able to adequately represent the spectrum of solution behaviour from ideal to highly non-ideal systems (54-56). The model contains as many curve-fitting parameters (usually 3) as is necessary to accurately describe the actual measured data.

The model was used to calculate multiple solubility maxima and also solute solubility in mixed solvents at various temperatures (57). The model was also used to correlate other physico-chemical properties (*PCP*) in mixed solvent systems; including the electrophoretic mobility of analytes in mixed solvent electrolyte systems (58-60), the instability rate constants in binary solvent systems (61), the acid dissociation constants in water-organic solvent mixtures at a fixed and various temperatures (52, 62), the capacity factor of analytes in HPLC (63), the dielectric constant (30),

surface tension (64), viscosity (65), density (66), solvatochromic parameter (67), refractive index (68) and ultrasound velocity (69) in the solvent mixtures. Theoretical basis of the model for describing the chemical potential of solutes dissolved in mixed solvents (19) and the acid dissociation constants in aqueous-organic mixtures (52) have been provided in earlier papers. The constants of the Jouyban-Acree model represent differences in the various solute-solvent and solvent-solvent interactions in the mixture (19). Therefore, the model should be able to calculate any *PCP* in mixed solvents, which is a function of solute-solvent and/or solvent-solvent interactions. The general form of the Jouyban-Acree model is:

$$\log PCP_{m,T} = f_1 \log PCP_{1,T} + f_2 \log PCP_{2,T} + f_1 f_2 \sum_{i=0}^{2} \frac{J_i (f_1 - f_2)^i}{T}$$

(37)

where $PCP_{m,T}$, $PCP_{1,T}$ and $PCP_{2,T}$ are the numerical values of the physico-chemical property of the mixture and solvents *1* and *2* at temperature *T*, respectively, f_1 and f_2 are the volume (weight or mole) fractions of solvents *1* and *2* in the mixture and J_i represent the model constants.

The model could be extended for representation of the *PCPs* in ternary solvents as:

$$\log PCP_{m,T} = f_1 \log PCP_{1,T} + f_2 \log PCP_{2,T} + f_3 \log PCP_{3,T}$$

$$+ f_1 f_2 \sum_{i=0}^{2} \frac{J_i (f_1 - f_2)^i}{T} + f_1 f_3 \sum_{i=0}^{2} \frac{J_i^{'} (f_1 - f_3)^i}{T}$$

$$+ f_2 f_3 \sum_{i=0}^{2} \frac{J_i^{''} (f_2 - f_3)^i}{T}$$

(38)

where subscript *3* is the solvent *3* characteristics, $J_i^{'}$ and $J_i^{''}$ are the sub-binary model constants.

The main limitations of the Jouyban-Acree model for predicting drug solubilities in water-cosolvent mixtures are: a) it requires two data points of solubilities in mono-solvent systems, and b) numerical values of the model constants. To overcome the first limitation, the solubility prediction methods in mono-solvent system should be improved. A number of articles reviewed the recent progresses in this field, especially with the aqueous solubility prediction methods (70-74). To address the second limitation, a number of solutions were examined:

i) the model constants could be obtained using solubility of structurally related drugs in a given water-cosolvent system, and then predict the un-measured solubility of the related drugs where the expected *MPD* was ~ 17 % (27).

ii) the model constants could be calculated using a minimum number of experimental data points, i.e. three data points, and then predict the solubilities at the rest of solvent compositions where the expected prediction *MPD* was < 15 % (75).

iii) the trained versions of the Jouyban-Acree models could be employed for solubility prediction of drugs in the aqueous mixtures of dioxane (76), ethanol (35), polyethylene glycol 400 (77) and propylene glycol (78) at various temperatures and the expected *MPDs* were ~ 27 %, ~ 48 %, ~ 40 % and ~ 24 %, respectively. Table 3 listed the numerical values of the Jouyban-Acree model constants for the 4 cosolvents studied. Further experimental data sets are required to train similar models for the other cosolvents.

iv) in the trained versions of the Jouyban-Acree model, we assumed the extent of the solute-solvent interactions are the same, however, it is not the case since various solutes possess different functional groups leading to various extent of the solute-solvent interactions. To cover this point, the deviated solubilities from the trained versions of the Jouyban-Acree model were correlated using QSPR models which resulted in the reduced *MPD* values. The *MPD* values for dioxane, ethanol, polyethylene glycol 400 and propylene glycol were 18, 33, 38 and 16 %, respectively (79, 80).

v) a generalized version of the Jouyban-Acree model was proposed using its combination with the Abraham parameters where the model constants of the Jouyban-Acree model were correlated with the functions of the Abraham solvent coefficients and the solute parameters as:

$$J_i = A_{0,i} + A_{1,i}(c_1 - c_2)^2 + A_{2,i} R_2 (r_1 - r_2)^2 + A_{3,i} \pi_2^H (s_1 - s_2)^2$$

$$+ A_{4,i} \sum \alpha_2^H (a_1 - a_2)^2 + A_{5,i} \sum \beta_2^H (b_1 - b_2)^2 + A_{6,i} V_X (v_1 - v_2)^2$$

(39)

where *A* terms were the model constants (81), *c*, *r*, *s*, *a*, *b* and *v* are the solvents coefficients, subscripts *1* and *2* denote cosolvent and water, respectively,

R_2 is the excess molar refraction, π_2^H is dipolarity/polarizability of the solute, $\sum \alpha_2^H$ denotes the solute's hydrogen-bond acidity, $\sum \beta_2^H$ stands for the solute's hydrogen-bond basicity and V_X is the McGowan volume of the solute in unit of $0.01(cm^3/mole)$. The J_i (i.e., J_0, J_1 and J_2) terms of the studied solubility data sets were regressed against $(c_1 - c_2)^2$, $R_2(r_1 - r_2)^2$, $\pi_2^H(s_1 - s_2)^2$, $\sum \alpha_2^H(a_1 - a_2)^2$, $\sum \beta_2^H(b_1 - b_2)^2$ and $V_X(v_1 - v_2)^2$ values to compute $A_{j,i}$ terms, then the $A_{j,i}$ terms were replaced in equation (37) and the solubility of drugs in binary solvents were predicted employing experimental values of X_1 and X_2. The applicability of the generalized model was evaluated employing 30 data sets including various cosolvents and the expected *MPD* for the predicted solubilities was ~ 18.5 % (81). Using this version, the only required data is the solubility in neat water and cosolvent systems. Table 4 listed the numerical values of the model constants of equation (39).

In practice, when the binary solvents are not able to dissolve the desired amount of a drug in a given volume, ternary solvents are often used. The applicability of the Jouyban-Acree model for calculating the solubility of drugs in ternary solvents was shown using solubility data of 19-nor 1α, 25-dihydroxyvitamin D_2 in water-ethanol-propylene glycol mixtures where the *MPD* was ~ 17 % (82). To provide a predictive model for ternary solvents based on solubility data in binary solvents, the sub-binary constants of the Jouyban-Acree model were used to predict the solubility of paracetamol in ternary solvents where the *MPD* was < 10 % (83). Using a minimum number of solubility data of salicylic acid in aqueous mixtures of ethanol and propylene glycol and also ethanol-propylene glycol mixtures (3 data points from each binary system), a trained version of the Jouyban-Acree model was presented to predict the solubility of salicylic acid in binary and ternary solvents where the *MPD* was ~ 7 % (84). These capabilities should be further investigated employing more experimental data sets.

The Jouyban-Acree model has theoretical justifications (19), showed the most accurate correlations among other cosolvency models (56) and capable of predicting the solubility data in mixed solvents at various temperatures employing

the solubility data in neat solvents (35, 76-78). It is also capable of modeling other physico-chemical properties in mixed solvents and we believe that it is the best cosolvency model both from correlative and predictive capabilities viewpoints and promises better applicability in the pharmaceutical applications.

In the latest work from our group, a new definition of the solubilization power of a cosolvent was presented. Based on the new definition the solubilization power of a cosolvent is defined as

$$\omega = \left(\log \frac{X_{m, max}}{X_2} \right)$$

in which the $X_{m, max}$ (maximum solubility in binary solvent) could be calculated using the Jouyban-Acree model employing the experimental values of X_1 and X_2 for a drug of interest. The more ω value for a cosolvent is the more solubilization power of the cosolvent which means a desired amount of a drug could be dissolved at the minimum concentration of the cosolvent (85). The concentration of the cosolvent in the pharmaceutical formulations is an important factor from both toxicological and economical points of view.

Table 3. Numerical values of the Jouyban-Acree model (J_0, J_1 and J_2) for commonly studied cosolvents, the number of data sets (*NDS*) employed in the training process of the model and the references.

Cosolvent	J_0	J_1	J_2	NDS	Reference
Dioxane	958.44	509.45	867.44	36	76
Ethanol	724.21	485.17	194.41	26	35
Polyethylene glycol 400	394.82	-355.28	388.89	79	77
Propylene glycol	37.03	319.49	-	27	78

THE MODIFIED WILSON MODEL

The modified Wilson model was derived from an expression for the excess free energy of mixing of non-electrolyte solutions based on the Flory-Huggins theory for athermal mixtures. Comor and Kopecni (86) modified the Wilson model and one can estimate solute solubility in binary solvent mixtures from measured values in the pure solvent and cosolvent and also excess Gibbs free energies for the binary solvent mixture (18). The modified Wilson model is:

$$\log\left(\frac{X_s^i}{X_m}\right) = 1 - \frac{f_1\left[1 - \log\left(\frac{X_s^i}{X_1}\right)\right]}{f_1 + f_2\Lambda_{12}^{adj}} - \frac{f_2\left[1 - \log\left(\frac{X_s^i}{X_2}\right)\right]}{f_1\Lambda_{21}^{adj} + f_2}$$

(40)

By adjustment of this model it was shown that a simplified form of the modified Wilson model (23) is able to calculate solute solubility in water-cosolvent mixtures more accurate than modified.

Wilson model, although this simplification was not successful in the case of solubility prediction in non-aqueous binary solvents (18). Thus the simplified form is:

$$-\log X_m = 1 - \frac{f_1(1 + \log X_1)}{f_1 + f_2\lambda_{12}^{adj}} - \frac{f_2(1 + \log X_2)}{f_1\lambda_{12}^{adj} + f_2}$$

(41)

where Λ_{12}^{adj}, Λ_{21}^{adj}, λ_{12}^{adj} and λ_{21}^{adj} are the adjustable parameters of the model which can be evaluated via a nonlinear least squares analysis or by developing a simple computer program which calculates the solute solubility at each composition of the cosolvent and employs pre-selected values for the adjustable parameters (23). By using equation (41), there is no need to determine the ideal mole fraction solubility of the solute, which is one of the advantages for equation (41). Nevertheless equations (40) and (41) are mathematically equivalent with the log-linear model, for which the adjustable parameters equal unity. The modified Wilson model is a correlative model.

THE GENERAL SINGLE MODEL

As a polynomial equation the general single model, GSM, (22) was derived from theoretically based cosolvency models, i.e. the excess free energy and the Jouyban-Acree models, by algebraic manipulations. It has already been used as an empirical equation to correlate solute solubility in the pharmaceutical literature (28, 87, 88). GSM is expressed as a single power series of the solute free cosolvent fractions by equation (42):

$$\log X_m = K_0 + K_1 f_1 + K_2 f_1^2 + K_3 f_1^3 + \cdots$$

(42)

where $K_0 - K_3$ denote the model constants, which are calculated using least squares analysis. Full details of the derivation was provided in the reference (22).

THE QSPR MODEL OF RYTTING *et al.*

Rytting *et al.* (89) determined the solubility of 122 solutes in water, polyethylene glycol 400 (PEG 400), and three binary mixtures containing 25, 50 and 75 % of PEG 400 and treated the solubility in binary solvents as separate systems. These models are slightly different from other cosolvency models which treated the binary solvents as a continous system. The general form of the Rytting model is:

$$\log X = c_0 + c_1 MW + c_2 V_s + c_3 RB + c_4 HBA$$
$$+ c_5 HBD + c_6 RG + c_7 D_s$$

(43)

where X is the molar solubility of the solute, MW is the molecular weight (g/mol), V_s is the molecular volume (Å^3), RB is the number of rotatable bonds, HBA is the number of hydrogen-bond acceptors, HBD is the number of hydrogen-bond donors, RG is the radius of gyration (Å), D_s is the molecular density ($\frac{MW}{V_s}$) and $c_0 - c_7$ are the model constants.

Table 4. The numerical values of the QSPR models (equation (39)) for predicting the model constants of the Jouyban-Acree model taken from a reference (81).

	A_0	A_1	A_2	A_3	A_4	A_5	A_6
J_0	2113.119	-1093.783	3380.661	-13.865	-4.921	-5.659	15.250
J_1	-2001.561	1142.780	-2735.160	-38.541	13.176	0.811	38.508
J_2	1474.963	-1507.479	4421.302	17.981	-21.196	6.595	-13.386

The investigated solutes were divided into 84 data as training set and 38 data as test set. The models (for each solvent composition) were trained using 84 data points and the prediction capability of the models were not satisfactory (89). To provide better models, genetic algorithm was used to divide the training set into two subsets 1 and 2 based on the similarities of the molecular descriptors. Two sets of the model constants ($c_0 - c_7$) were reported and the solubilities of the test compounds were predicted using the most similar model of 1 or 2 to the descriptors of the test compound. The similarity between compounds were determined using d_{ij} term defined by:

$$d_{ij} = \sum (X_{ik} - X_{jk})^2 \qquad k = 1-7$$

(44)

where X_{i1}, X_{i2}, … and X_{i7} are the seven desciptors of compound i and X_{j1}, X_{j2}, … and X_{j7} are the seven desciptors of compound j and d_{ij} is the elucidation distance between compounds i and j. The smallest d_{ij} was found between test set compound A and training set compound B, means that solubility prediction of compound A using trained model by set B will produce the best predictions.

Rytting *et al.* (89) evaluated the accuracy of their models by computing the residual ranges in log unit for the test set, sorted in five ranges, i.e. $<\pm$ 0.5, $\pm$ 0.5-1.0, $\pm$ 1.0-1.5, $\pm$ 1.5-2.0 and $>\pm$2.0. The results were also compared with those of the log-linear model of Yalkowsky. Figure 2 showed the relative frequencies of the residuals of the QSPR model of Rytting *et al.* and log-linear model of Yalkowsky for three water- PEG 400 compositions.

The relative frequencies of both models are similar, however, it should be noted that the log-linear model employs the aqueous solubility data of each solute as input value and the QSPR model treated each solvent composition as a separate system. In addition, the required computations are more complicated when compared with the straigthforward calculations of the log-linear model.

THE ARTIFICIAL NEURAL NETWORK MODEL

An attempt was made using artificial neural network (ANN) models and different numerical analyses which are demanded in the pharmaceutical applications (90). The optimized topology of the ANN was 6-5-1 and the network was used f_1, f_2, $-\log X_1$, $-\log X_2$, δ_1 and δ_2 as input variables and $-\log X_m$ was its output. The accuracy of the ANN method for computing solubility data of drugs was checked by calculating the mean percentage deviation (*MPD*) and individual percentage deviation (*IPD*) for 35 data sets of various drugs in 8 cosolvent systems. The results were compared with the best muliple linear regression (MLR) model (i.e. the Jouyban-Acree model).

In numerical analysis I, all data points from each set was used to train the models and the back-calculated solubilities were used to compute the *MPD* and *IPD* values. The overall *MPDs* were 0.9 and 5.6 % for ANN and MLR models, respectively. In analysis II, five data points from each data set were used as trainig set and the solubility at other solvent compositions were predicted. The overall *MPDs* for ANN and MLR models were 9.0 and 11.3 %, respectively. A single ANN model (analysis III) was trained using all data points of 35 data sets and the solubilities were back-calculated where the ovearll *MPD* was 24.8 %. In the numerical analayis IV, solubility data sets with odd numbers were used as training set for ANN and the sets with even numbers were used as prediction set and the overall *MPD* for predicted solubilities was 56.0 %.

The correlation ability of ANN and MLR models for the solubility of various drugs in aqueous mixtures of a given cosolvent were investigated (analysis V), where the overall *MPDs* were 2.0 and 20.4, respectively for ANN and MLR models. In numerical analysis VI, the correlation ability of the models for structurally related drugs in a given water-cosolvent mixtures were studied. The ANN model produced the overall *MPD* of 4.7 % in comparison with the 18.4 % of MLR model. In the last numerical analysis (VII), the correlation ability of a given drug in various water-cosolvent systems was investigated. The ANN model produced overall *MPD* of 3.4 %, whereas the corresponding value for the MLR model was 67.2 %.
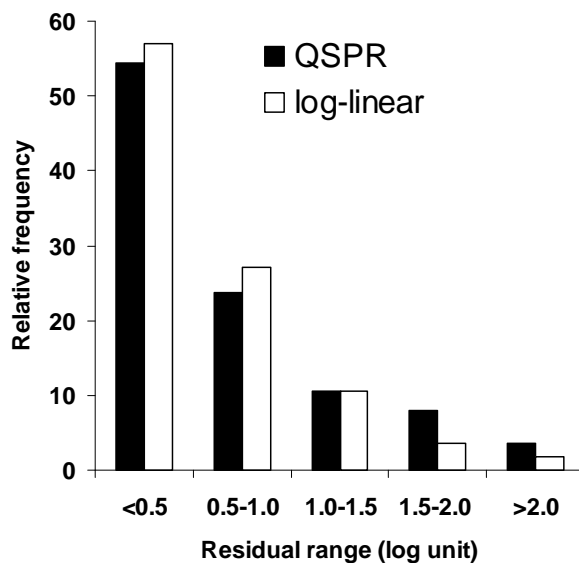
Figure 2. The relative frequencies of the residuals of the QSPR and log-linear models sorted in five groups (Figure is drawn in this work and its raw data taken from the original reference (89)).
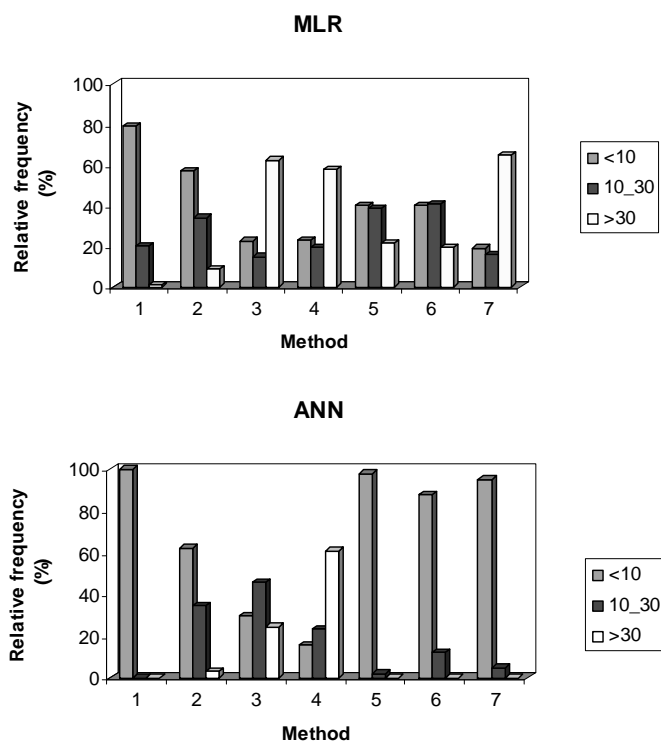

Figure 3. Relative frequency of individual percentage deviations (*IPD*) for numerical methods 1-7 using MLR and ANN models (taken from a reference (90))

The mean differences of the overall *MPDs* of the ANN and MLR models were statitstically significant in all numerical analyses except than analysis IV where there was no significant difference between two models. The ANN method produced better IPD distribution in comparison with the MLR model as shown in Figure 3 (90).

THE UNIFIED COSOLVENCY MODEL

As discussed in this review, several mathematical equations were presented to calculate the solute solubility in mixed solvents. The models are derived using different assumptions and are even based on empirical procedures. The solubility in mixed solvents is a single phenomenon and it is expected to mathematically represent the system using a single equation. In order to show that this is the case for the cosolvency models, a unified cosolvency model was derived which considers all available cosolvency models. More details of derivation procedure was presented in a previous paper (91). For this, we can summarise all cosolvency models as a power series of volume fraction of the cosolvent, GSM model. The main differences in the cosolvency models are that the accuracy and predictability of most of models are different from each other. This is the case because the models employed a different arrangement of independent variables and also different number of model constants (91).

THE COSMO-RS MODEL

The conductor-like screening model for real solvents (COSMO-RS) is a predictive model which integrates the concepts of quantum theory, dilectric contiuum models and surface interactions. Ikeda *et al.* (92) predicted the solubility of 15 drugs in water, ethanol, acetone and chloroform and compared the predicted solubilities with the experimental values by computing root mean square error (*RMSE*) which its overall value was 0.64 in log unit. The authors also tested the prediction capability of the COSMO-RS model on oxolinic acid solubility data in water-ethanol and sulfadiazine solubility data in water-dioxane mixtures where the *RMSE* values were 2.0 and 1.6 log unit, respectively for oxolinic acid and sulfadiazine data (92). In addition to these deviations, the model applications require relatively sophisticated computations which are not favoured methods for most of pharmaceutical scientists.

THE NEW MODELS PROPOSED BY YALKOWSKY'S GROUP

A new model was proposed by Machatha *et al.* (93) which is claimed to be a better predictor of solubility of drugs in water-ethanol mixtures when compared with the log-linear model. The model is:

$$\log X_m = \frac{\log X_2 + a \cdot f_1}{1 + b \cdot f_1 + c \cdot f_1^2} \qquad (45)$$

where *a*, *b* and *c* are the model constants. The authors compared the accuracy of equation (45) with that of GSM model (the third order polynomial of cosolvent fractions) employing solubility of 51 compounds in water-ethanol mixtures. The accuracy criteria used in the comparison were root mean square error (*RMSE*) and average absolute error (*AAE*), where the average *RMSE* values for equation (45) and the third order polynomial were 0.035 and 0.049, respectively. The corresponding *RMSE* for the flactuation model of Ruckenstein was 0.064 (93).

A bilinear function that accounts for the disparity between log-linear and parabolic models was also presented by Machatha and Yalkowsky (94) and its accuracy was checked using *RMSE* criterion by emloying the solubility data of 52 sets in water-ethanol mixtures. The proposed model for water-ethanol mixtures is:

$$\log\left(\frac{X_m}{X_2}\right) = \sigma_A \cdot f_1 + \frac{(\sigma_A - \sigma_B)f_1}{1 + 10^{-3.6(f_1-1)}} \qquad (46)$$

where σ_A is the slope of the ascending part of the solubility profile and σ_B is the slope of the descending section of the profile. The *RMSE* values for the investigated data sets were reported in Table 5 along with the *RMSE* of the equation (45), the third order polynomial, the second order polynomial, bilinear (equation (46)) and log-linear (equation (5)) models collected from the literature (94, 95). It should be noted that the average *RMSE* of the bilinear model was more than that of equation (45).

ACCURACY CRITERIA USED IN THE LITERATURE

The accuracy of the predicted solubilities using the cosolvency models should be evaluated to recognise the differences between predictive and less predictive models and also to provide the expected error range for a preditive model. Various accuracy criteria using various terminologies were used in the literature which is briefly reviwed in this work. Yalkowsky, the pioneer of cosolvency modeling investigations, and his co-workers (93, 94) often used the root mean square error (RMSE) and defined it as:

$$RMSE = \sqrt{\frac{\sum \left(Observed - \Pr edicted\right)^2}{N}} \quad (47)$$

where N is the number of predicted solubility data points. The next criteria used by Yalkowsky's group (93, 95) is the average absolute error (AAE) which is defined as:

$$AAE = \frac{\sum \left|Observed - \Pr edicted\right|}{N} \quad (48)$$

The RMSE and AAE values could be defined in logarithmic ($\log X_m$ or $\ln X_m$) or arithmatic (X_m) scales and to compare its reported values in different papers, the scale should be kept in mind.

Table 5. The root mean square error (RMSE) of various equations for calculating solute solubility in water-ethanol mixtures taken from the references (94-95).

Solute	N^a	Equation (45)	Equation (42) up to power 3	Equation (42) up to power 2	Equation (46)	Equation (5)
Acetanilide	13	0.021	0.027	0.070	0.041	0.248
Alanine	10	0.010	0.055	0.195	0.050	0.982
Alprazolam	9	0.022	0.035	0.103	0.063	0.365
p-Aminobenzoic acid	6	0.023	0.036	0.044	0.056	0.256
Aminocaproic acid	10	0.029	0.105	0.271	0.047	0.988
Amino-isobutyric acid	5	0.004	0.007	0.025	0.017	0.179
Amino-n-butyric acid	6	0.005	0.044	0.140	0.017	0.574
Anthracene	11	0.061	0.091	0.129	0.123	0.267
Asparagine	5	0.010	0.019	0.128	0.003	0.627
Aspartic acid	9	0.062	0.104	0.220	0.092	0.759
Barbital	11	0.018	0.014	0.061	0.035	0.214
Benzamide	14	0.019	0.012	0.061	0.031	0.324
Benzocaine	11	0.038	0.048	0.110	0.083	0.345
Benzoic acid	11	0.029	0.083	0.103	0.112	0.543
Biphenyl	11	0.069	0.092	0.152	0.204	0.200
Caffeine	6	0.026	0.032	0.135	0.000	0.582
Camphoric acid	12	0.047	0.052	0.069	0.201	0.414
Diazepam	11	0.054	0.057	0.122	0.086	0.513
Didanosine	11	0.050	0.062	0.081	0.075	0.456
beta-Estradiol	6	0.034	0.085	0.155	0.134	0.645
5-Ethylhydantoin	7	0.071	0.019	0.071	0.023	0.296
Formyl-aminobutyric acid	7	0.051	0.011	0.069	0.024	0.211
Formylglycine	9	0.008	0.027	0.074	0.020	0.288
Formylleucine	8	0.043	0.063	0.083	0.076	0.282
Furosemide	13	0.148	0.133	0.361	0.420	0.364
Glutamic acid	6	0.038	0.102	0.217	0.076	0.790
Glutamine	5	0.006	0.007	0.007	0.034	0.623
Glycine	10	0.016	0.072	0.158	0.035	0.148
Glycylglycine	7	0.016	0.074	0.213	0.025	0.854

Table 5 continued…

Histidine	8	0.018	0.033	0.087	0.005	0.275
Hydantoic acid	6	0.016	0.020	0.094	0.014	0.380
Hydantoin	7	0.020	0.027	0.093	0.018	0.356
Ibuprofen	8	0.120	0.123	0.249	0.197	0.462
Indomethacine	10	0.058	0.083	0.172	0.147	0.419
Leucine	5	0.022	0.014	-	-	-
Metharbital	11	0.018	0.028	0.078	0.053	0.279
Methylhydantoic acid	6	0.020	0.020	0.095	0.023	0.294
Naphthalene	6	-	-	0.000	0.000	0.000
Norleucine	10	0.035	0.055	0.137	0.057	0.316
Oxolinic acid	11	0.049	0.063	0.065	0.132	0.147
Paracetamol	13	0.083	0.034	0.057	0.053	0.349
Phenobarbital	12	0.017	0.015	0.092	0.087	0.488
Phenylalanine	8	0.018	0.040	0.097	0.052	0.457
Phenytoin	11	0.046	0.053	0.122	0.090	0.419
Salicylic acid	6	0.007	0.067	0.129	0.106	0.258
Strychnine	7	0.038	0.054	0.119	0.055	0.508
Tartaric acid	12	0.002	0.005	0.017	0.005	0.092
Theophylline	10	-	-	0.080	0.027	0.364
Triglycine	7	0.057	0.079	0.277	0.025	1.010
Tryptophan	8	0.042	0.019	0.097	0.050	0.372
DL-Valine	7	0.026	0.058	0.146	0.043	0.461
Zalcitabine	11	0.022	0.026	0.052	0.056	0.396
Ziduvudine	11	0.022	0.024	0.148	0.032	0.257
Overall		0.035	0.050	0.120	0.069	0.423

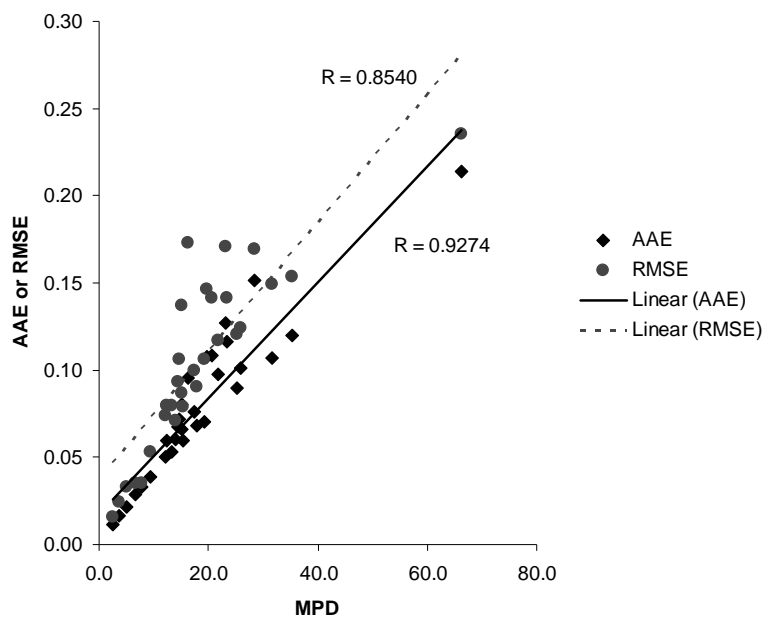[a]N is the number of data points in each set.

Figure 4. Correlations between various error criteria; average absolute error (*AAE*), root mean square error (*RMSE*) with mean precentage deviation (*MPD*) (Figure is reproduced from a reference (81)).

The *%Dev.* was used by Acree and his colleagues and defined as:

$$\%Dev. = \left(\frac{100}{N}\right)\sum\left|\log\left(\frac{Pr\,edicted}{Observed}\right)\right| \qquad (49)$$

The mean squared deviation (*MSD*) was also used in the literature:

$$MSD = \sqrt{\frac{\left(X_m^{Calculated} - X_m^{Observed}\right)^2}{N-1-p}} \qquad (50)$$

where *p* is the number of the parameters of the model (96).

The mean percentage deviation (*MPD*) was used by our group (23, 76-81, 83, 84) and is defined as:

$$MPD = \frac{100}{N}\sum_1^N\left(\frac{\left|X_m^{Calculated} - X_m^{Observed}\right|}{X_m^{Observed}}\right) \qquad (51)$$

The same definition was also used in some reports using various terminologies, such as percent mean error (27, 48, 57), average percentage deviation (75, 117, 118) and percent deviation (56).

There are good correlations between various error criteria as it has been shown Figure 4 (81), however, the data scattered around the line and in some cases significant deviation was observed when a model assessed using *RMSE* or *MPD* criteria. This is not an important parameter in comparing the accuracy of the models using a criterion, however, the experimental investigator needs to know the range of prediction error from a practical point of view and compare the accuracy of the experimentally determined data and their relative standard deviations with the prediction error. From this point of view, the *MPD* definition may be the best error criterion.

ACCEPTABLE *MPD* RANGE IN SOLUBILITY CALCULATIONS

There is a difference between "correlation" and "prediction" terms which are often used in the literature. In "correlation", the experimental data points were fitted to the model and the models fitness capability could be assessed using the results of correlation studies. When the capability of a model was shown using the "correlation" studies, the model could be used for predicting unmeasured experimental data. This capability could be evaluated by dividing the exsiting experimental data into training and test sets. The predicted data of test set by the trained model was compared with the experimentally obtained data using an accuracy criterion, like *MPD* or *RMSE* etc. This process could be named as "prediction".

The solubility calculation can be studied from the calculation error point of view. In order to provide a comprehensive approach to evaluate the process of solubility prediction, these calculations have been divided into 3 groups:

1) Pure predictive calculations such as group contribution methods; these types of calculations have been widely employed to predict the aqueous solubility of pharmaceutical and environmentally interesting compounds. A method using the aqueous functional group activity coefficients (AQUAFAC) have been presented for predicting aqueous solubility of 168 nitrogen containing compounds including some pharmaceuticals (97). The results are reported as the absolute average error (*AAE*) on a logarithmic scale as 0.42. However, this information is not in common format, and in order to provide a meaningful error term, which is comparable with RSD values, the deviations between predicted and experimental values, was calculated using equation (51). The calculated *MPD* value based on tabulated aqueous solubility data of 165 compounds (97) was 2108 % (98). The individual percentage deviations span between 0 % for nonylamine to 308219 % for dipropalin (its logarithm of aqueous solubility ($\log X_2$) is –2.966 and the predicted $\log X_2$ is -6.455). The *MPD* of AQUAFAC is still very high (*MPD*=284 %) when 5 outliers which produced very high errors were taken out (98).

2) Semi-predictive calculations such as the general solubility equation (GSE) proposed by Yalkowsky and Valvani (99); The bases of these calculations were to employ a training set to compute the model constants and then from this to predict the other data points. In a comparative study reported by Huuskonen *et al.* (100), three multiple linear regression methods (including GSE) and an artificial neural network (ANN) model have been trained using 675 aqueous solubility data points taken from databases. In this data, the prediction capabilities of the models have been tested using aqueous solubility data of 38 pharmaceuticals. The *MPD* values for multiple linear regression models were 2466 ± 12877, 117 ± 125 and 181 ± 327 % and the corresponding value for AAN model was

154 ± 193 (98). Employing another version of the equations, Abraham and Le (101) proposed an amended solvation energy relationship to calculate the aqueous solubility of solutes. They employed a training set, which includes the aqueous solubility of 659 compounds and then from this predicted the aqueous solubility of 65 compounds. The reported *AAE* for the 65 predicted solubility values on a logarithmic scale was 0.50. The corresponding *MPD* value obtained was 134 ± 373 % (98). These *MPD* values are so large as to be meaningless and therefore it is suggested that such models would not be of use in the industry. However even with such enormous errors, both types of the solubility calculations are highly in demand in the pharmaceutical area, where fast, reliable and generally applicable methods are requested for prediction of aqueous solubility of new drugs before a promising drug candidate has been progressed through development process. From this information it is clear that current knowledge and expertise are insufficient to guarantee fully the reliability of the predictions. Thus, they are not acceptable as a basic procedure prior to developing an optimised process. But they do serve as a rough guide for further testing and for priority setting of parameter range.

3)   Correlative calculations such as simple least square methods; These methods provide a mean to screen the experimental data to detect the possible outliers where re-determination is required. In addition the correlative trained models can be employed to predict undetermined data. As an example, the GSE model is proposed for correlating aqueous solubility of different compounds in water. Its correlation ability and that of a revised form of GSE have been studied by Jain and Yalkowsky (102). However again although the *MPD* values for original and revised forms of GSE have been calculated employing aqueous solubility of 582 nonelectrolytes, the *MPDs* obtained are 2881 ± 30716 and 1417 ± 15393 %, respectively and therefore are meaningless (98). The reported error percentage for the correlative form of an amended solvation energy relationship employing 659 data was 0.408 (101) and the corresponding *MPD* value was 244 ± 850 % (98).

A summary of different types of solubility calculations has been shown in Table 6. The deviations from experimental solubilities have been presented by *MPD*, and where there was no possibility to calculate the *MPD*, the original reported criteria collected form the papers have been employed.

In trying to explain why prediction errors are so high, a number of sources can be discussed. In addition to the errors related to the nature of the models and calculation procedure, the quality of solubility data is a possible contribution to error sources in solubility data modelling. Experimental solubility data for a given solute can vary from laboratory to laboratory. As examples, a couple of solubility data collected from the literature are shown in Table 7. The compilation of solubility data given in Table 7 clearly illustrates that there can be a significant amount of variability in experimental solubilities. In an interesting work by Kishi and Hashimoto (112), solubility data of antharacene and fluoranthene reported by 17 different laboratories using a standard method by the environmental agency of Japan have been summarized. The results showed that even when all variables were kept constant, inter-laboratory difference can still be very significant. They showed that the mean solubilities of anthracene and fluoranthene span 0.17 and 0.36 log unit and *MPD* value is 51 % (112). It is obvious that the ranges of individual solubilities are even greater where the range of logarithm of mole/L solubility for anthracene was –7.08 to –5.23. In addition to the variation of results for different laboratories, the RSD values for the repeated experiments from the same laboratory are significantly high. As examples, the reported RSD values are up to 9.2 % (20), 4.4 % (113) and 10 % (114). The possible reasons for such differences in solubilities arise from: 1) solute purity, 2) equilibration time, 3) temperature, 4) analysis method, 5) laboratory technique (115), 6) typographical error, 7) polymorphism (116) and 8) enantiomeric forms (111).

From a computationl chemist's point of view, the ideal model would have a *MPD* value close to zero. However, this is impossible because of small variations in experimental results (uncertainty) that appear in the training process of the model. In practice the experimental chemists look for a model, which enables them to calculate the solubility with *MPD* lower than the RSD values which are obtained from repeated experiments.

As shown in Table 6 the error levels for aqueous solubility calculations are between 117 to 2500 %. Dickhut *et al.* (103) proposed that a mixed solvent solubility estimation method is considered acceptable if solubility calculations are, on the

average, within 30 % of the experimentally determined values. This has also been confirmed by other research groups (36, 37). From our point of view, any model providing more accurate correlations/predictions in comparison with the previous models could be considered as a step forward and a single model could not produce the most accurate results for all data sets studied.

COMPARISON OF THE ACCURACY OF THE MODELS

An extensive comparison between the accuracy of various cosolvency models with different curve-fitting parameters for correlating the experimental data was reported employing 30 solubility data sets of pharmaceutical compounds in 7 cosolvents (56).

Table 6. The reported deviations for solubility calculations collected from the literature

System[a]	Model	Reference	Correlative calculations		Predictive calculations	
			N[b]	Deviation	N[b]	Deviation
I	UNIFAC	103	5 to 17	MPD[c]=55	-	-
I	14	103	5 to 17	MPD=36	-	-
I	19	103	5 to 17	MPD=49	-	-
I	45	93	11	RMSE[d]=0.049	-	-
I	42	93	11	RMSE=0.063	-	-
I	Ruckenstein and Shulgin	93	11	RMSE=0.064	-	-
I	42 (up to power 2)	95	-	-	460	RMSE=0.120
I	5	35	294	AAE=0.48	-	-
I, II	22	117	12 to 13	MPD=6.9	7 to 8	MPD=7.3
I, II	26	117	12 to 13	MPD=11.4	7 to 8	MPD=14.6
I, II	36	117	12 to 13	MPD=4.9	7 to 8	MPD=7.9
I, II	42 (up to power 4)	117	12 to 13	MPD=7.5	7 to 8	MPD=8.0
I, II	Mean predicted solubility	117	12 to 13	MPD=5.9	7 to 8	MPD=6.4
I, II	36	90	11 to 21	MPD=5.56	6 to 16	MPD=10.33
I, II	ANN	90	11 to 21	MPD=0.90	6 to 16	MPD=9.04
II	22	117	-	-	6 to 8	MPD=8.7
II	26	117	-	-	6 to 8	MPD=13.3
II	36	117	-	-	6 to 8	MPD=7.2
II	42 (up to power 4)	117	-	-	6 to 8	MPD=7.4
II	Mean predicted solubility	117	-	-	6 to 8	MPD=7.6
II	42 (up to power 3)	75	-	-	6 to 16	MPD=13.1
II	26	75	-	-	6 to 16	MPD=14.2
II	36	75	-	-	6 to 16	MPD=12.5
II	41	75	-	-	6 to 16	MPD=15.0
II	Mean predicted solubility	75	-	-	6 to 16	MPD=10.6
II	45	93	-	-	455	RMSE=0.035
II	42	93	-	-	455	RMSE=0.050
II	5	95	-	-	460	RMSE=0.423
II	45	95	-	-	460	RMSE=0.069
II	37	35	294	AAE[e]=0.19	-	-
II	37, 39	81	-	-	467	MPD=18.5
II	5	81	-	-	467	MPD=76.7
III	36	90	496	MPD=90.42	-	-
III	ANN	90	496	MPD=24.76	-	-
IV	36	90	-	-	236	MPD=81.10
IV	ANN	90	-	-	236	MPD=55.97

Table 6 continued…

V	36	90	454	MPD=20.36	-	-
V	ANN	90	454	MPD=2.02	-	-
VI	14 modified	26	142	MPD=35	-	-
VI	36	27	142	MPD=18	-	-
VI	36	90	278	MPD=18.37	-	-
VI	ANN	90	278	MPD=4.70	-	-
VII	36	90	120	MPD=67.19	-	-
VII	ANN	90	120	MPD=3.36	-	-
VIII	Least square	104	331	MSD=0.30	17	MSD[f]=0.34
VIII	AQUAFAC	96	873	MSD=0.56	97	MSD=0.56
VIII	GSE	96	873	MSD=0.80	97	MSD=0.80
VIII	ANN	105	123	MSD=0.22	13	MSD=0.23
VIII	Least square	105	123	MSD=0.28	13	MSD=0.28
VIII	AQUAFAC	106	133	MSD=0.33	25	MSD=0.324
VIII	GSE	106	-	-	25	MSD=0.423
VIII	Mobile order	107	-	-	531	MSD=0.37
VIII	Least square	100	675	-	38	MPD=2466
VIII	Least square	100	675	-	38	MPD=117
VIII	Least square	100	675	-	38	MPD=181
VIII	ANN	100	675	-	38	MPD=154
VIII	GSE	102	582	MPD=2881	-	-
VIII	GSE-revised form	102	582	MPD=1417	-	-

[a] System, I: solubility of a given solute in binary aqueous solvent mixture, II: solubility of a given solute in binary aqueous solvent mixture using the trained model employing a minimum number of experimental data points III: correlation of solubility of various solutes in different water-cosolvent mixtures, IV: prediction of solubility of various solutes in different water-cosolvent mixtures, V: correlation of solubility of various solutes in a given water-cosolvent mixture, VI: correlation of solubility of structurally related solutes in a given water-cosolvent mixture, VII: correlation of solubility of a given solute in different water-cosolvent mixtures and VIII: correlation of aqueous solubility of different solutes.

[b] N is the number of correlated/predicted data points in each set.

[c] MPD is the mean precentage deviation and is calculated by equation (51).

[d] $RMSE$ is the root mean square error (in log scale) and is calculated by equation (47).

[e] AAE is the average absolute error (in log scale) and is calculated by equation (48).

[f] MSD is the mean squared deviation and is calculated by equation (50).

Table 7. The differences between aqueous solubility of drugs from different research groups

No.	Solute	Reference 1	Solubility 1	Reference 2	Solubility 2	% Difference
1	Paracetamol	108	0.001	110	0.001915	191
2	Sulfadiazine	109	0.00000493	28	0.00000327	51
3	Sulfadimidine	109	0.0000274	26	0.00000302	804
4	Sulfapyridine	109	0.0000194	36	0.00000178	988
5	Ibuprofen	101	0.000174	111	0.00943 (Racemate)	98
6	Theophylline	101	0.0407	20	0.0340	17
7	Naphthalene	102	0.000251	20	0.000201	25

The cosolvency models possess different numbers of curve-fitting parameters and require various number of experimental data in the training process. As an example, the MPD of the general single model reduced from ~ 16 % to ~ 5 % by using 3 and 7 curve-fitting parameters, respectively (56). The correlation ability of equations (23), (26), (36), (41) and (42) was compared where the overall MPD values were 5.92, 6.33, 3.06, 7.77 and 5.06 %, respectively, and equation (36) was the best from

correlation ability point of view (56). The corresponding values for the *MPD* of the Ruckenstein-Shulgin model was 7.7 % (119). This model is derived from the Kirkwood-Buff theory of solutions and required sophisticated caculations and is not a favoured cosolvency model in pharmaceutical area and therefore is not considered in this review. This was also the case for a recent work of Mirmehrabi et al. (120).

The accuracy of the simplest cosolvency model; i.e. the log-linear model of Yalkowsky, was compared with that of the gererally trained version of the Jouyban-Acree model employing 26 data sets in ethanol-water mixtures at various temperatures. The reported *AAEs* of the log-linear and Jouyban-Acree models were 0.48 ± 0.28 and 0.19 ± 0.13, respectively (35).

CONCLUSION

Solubility prediction in pharmaceutical area is still a challenging subject and requires further investigations from both experimental and computational points of view. The most commen method for increasing the low aqueous solubility of drugs is the addition of a miscible organic solvent, cosolvency. In addition to the experimental efforts for developing fast and reliable solubility determination methods and reporting new experimental solubility data in water-cosolvent mixtures, a number of cosolvency models have been proposed to represent the data. The available cosolvency models were reviewed along with a short discussion on the acceptable prediction error, the accuracy of the present models and also the accuracy criteria from the literature were discussed. From the present models, the log-linear model of Yalkowsky provides a good estimate of the drug solubilities in water-cosolvent mixtures employing aqueous solubility datum and the Jouyban-Acree model provides acceptable predictions to be used in the pharmaceutical industry. The latter model has theoretical justification and could also be used for modelling other phisico-chemical properties in mixed solvents and promises better applicabilities in pharmaceutical area.

ACKNOWLEDGMENT

REFERENCES

[1].   Yalkowsky S.H.; He Y. Handbook of Aqueous Solubility Data, CRC, Boca Raton, FL, USA, 2003.

[2].   Myrdal, P.B.; Yalkowsky, S.H. Solubilization of drugs. In: Swarbrick, J.; Boylan, J.C. (Eds.), Encyclopedia of Pharmaceutical Technology, Dekker, New York, Vol. 18: 161-217, 1998.

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

[4].   Spiegel AJ, Noseworthy MM. Use of nonaqueous solvents in parenteral products. J Pharm Sci, 52: 917-927, 1963.

[5].   Patel DM, Bernardo P, Cooper J, Forrester RB. Glycerine. In: Handbook of Pharmaceutical Excipients. American and Great British Pharmaceutical Societies, London, pp. 203-213, 1986.

[6].   Tsai PS, Lipper RA, Worthington HC. Propylene glycol. In: Handbook of Pharmaceutical Excipients. American and Great British Pharmaceutical Societies, London, pp. 241-242, 1986.

[7].   Golightly LK, Smolinkse SS, Bennett ML, Sunderland III EW, Rumack BH. Pharmaceutical excipients adverse effects associated with inactive ingredients in drug products (Part I). Med. Toxicol. 3: 128-165, 1988.

[8].   Rubino JT. Cosolvents and cosolvency, In: Swarbrick, J.; Boylan, J.C. (Eds.), Encyclopedia of Pharmaceutical Technology, Dekker, New York, Vol. 3:, 375-398, 1990.

[9].   Allen MJ, Borody TJ, Bugliosi TF, May GR, LaRusso NF, Thistle JL. Rapid dissolution of gallstones by methyl tert-butyl ether- preliminary observations. New Eng J Med, 312: 217-220, 1985.

[10].   Li A, Andren AW. Solubility of polychlorinated biphenyls in water/alcohol mixtures. 1. Experimental Data. Environ Sci Technol, 28: 47-52, 1994.

[11].   Sunderland VB, Watts DW. Alkaline ethanolysis of methyl 4-hydroxybenzoate and hydrolysis of methyl and ethyl 4-hydroxybenzoates in ethanol-water systems, Int J Pharm, 27: 1-15, 1985.

[12].   Chertkoff MJ, Martin AN. The solubility of benzoic acid in mixed solvents. J Am Pharm Assoc, 49: 444-447, 1960.

[13].   Paruta AN, Sciarrone BJ, Lordi NG. Solubility of salicylic acid as a function of dielectric constant. J Pharm Sci, 53: 1349-1353, 1964.

[14]. Adjei A, Newburger J, Martin A. Extended Hildebrand approach. Solubility of caffeine in dioxane-water mixtures. J Pharm Sci, 69: 659-661, 1980.

[15]. Yalkowsky, S.H.; Roseman, T. In: Yalkowsky, S.H. (Ed.), Solubilization of Drugs by Cosolvents. Marcel Dekker, New York, pp. 91-134, 1981.

[16]. Williams NA, Amidon GL. Excess free energy approach to the estimation of solubility in mixed solvent system. II. Ethanol-water mixtures. J Pharm Sci, 73: 14-18, 1984.

[17]. Ochsner AB, Belloto Jr, RJ, Sokoloski TD. Prediction of xanthine solubilities using statistical techniques. J Pharm Sci, 74: 132-135, 1985.

[18]. Acree Jr WE, McCargar JW, Zvaigzne AI, Teng IL. Mathematical representation of thermodynamic properties. Carbazole solubilities in binary alkane + dibutyl ether and alkane + tetrahydropyran solvent mixtures. Phys Chem Liq, 23: 27-35, 1991.

[19]. Acree Jr WE. Mathematical representation of thermodynamic properties. Part II. 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: 71-79, 1992.

[20]. Khossravi D, Connors KA. Solvent effect on chemical processes. I: Solubility of aromatic and heterocyclic compounds in binary aqueous-organic solvents. J Pharm Sci, 81: 371-379, 1992.

[21]. McHale MER, Pandey S, Acree Jr WE. Thermochemical Investigations of Hydrogen-Bonded Solutions: Part 12. Development of Expression for Predicting Solute Solubility in Binary Alcohol + Water Solvent Mixtures Based upon Mobile Order Theory. Phys Chem Liq, 33: 93-112, 1996b.

[22]. Barzegar-Jalali M, Jouyban-Gharamaleki A. A general model from theoretical cosolvency models. Int J Pharm, 152: 247-250, 1997.

[23]. Jouyban-Gharamaleki A. The Modified Wilson Model and Predicting Drug Solubility in Water-Cosolvent Mixtures. Chem Pharm Bull, 46: 1058-1061, 1998.

[24]. Fredenslund A, Jones RL, Prausnitz JM. Group contribution estimation of activity coefficients in nonideal liquid mixtures. AIChE J, 21: 1086-1099, 1977.

[25]. Acree Jr WE. Solubility of biphenyl in binary solvent mixtures. Int J Pharm, 18: 47-52, 1984.

[26]. Bustamante P, Escalera B, Martin A, Selles E. A modification of the extended Hildebrand approach to predict the solubility of structurally related drugs in solvent mixtures. J Pharm Pharmacol, 45: 253-257, 1993.

[27]. Jouyban-Gharamaleki A, Barzegar-Jalali M, Acree Jr WE. Solubility correlation of structurally related drugs in binary solvent mixtures. Int J Pharm, 166: 205-209, 1998.

[28]. Martin A, Wu PL, Adjei A, Lindstrom RE, Elworthy PH. Extended Hildebrand solubility approach and the log linear solubility equation. J Pharm Sci, 71: 849-856, 1982.

[29]. Yalkowsky SH, Amidon GL, Zografi G, Flynn GL. Solubility of nonelectrolytes in polar solvents III: Alkyl p-aminobenzoates in polar and mixed solvents. J Pharm Sci, 64: 48-52, 1975.

[30]. Jouyban A, Soltanpour Sh, Chan HK. A simple relationship between dielectric constant of mixed solvents with solvent composition and temperature. Int J Pharm, 269: 353-360. 2004.

[31]. Martin A, Paruta AN, Adjei A. Extended-Hildebrand solubility approach: methylxanthines in mixed solvents. J Pharm Sci, 70: 1115-1120, 1981.

[32]. Valvani SC, Yalkowsky SH, Roseman TJ. Solubility and partitioning IV: Aqueous solubility and octanol-water partition coefficients of liquid nonelectrolytes. J Pharm Sci, 70: 502-507, 1981.

[33]. Li A, Yalkowsky SH. Predicting cosolvency. 1. Solubility ratio and solute $logK_{ow}$. Ind Eng Chem Res, 37: 4470-4475, 1998.

[34]. Millard JF, Alvarez-Nunez FA, Yalkowsky SH. Solubilization by cosolvents. Establishment useful constants for the log-linear model. Int J Pharm, 245: 153-166, 2002.

[35]. Jouyban A, Acree Jr WE. In silico prediction of drug solubility in water-ethanol mixtures using Jouyban-Acree model. J Pharm Pharmaceut Sci, 9: 262-269, 2006.

[36]. Reillo A, Cordoba M, Escalera B, Selles E, Cordoba Jr, M. Prediction of sulfamethiazole solubility in dioxane-water mixtres. Pharmazie, 50: 472-475, 1995.

[37]. Beerbower A, Wu PL, Martin A. Expanded solubility parameter approach 1. Naphthalene and benzoic acid in individual solvents. J Pharm Sci, 73: 179-188, 1984.

[38]. Li A. Predicting cosolvency. 3. Evaluation of the extended log-linear. Ind Eng Chem Res, 40: 5029-5035, 2001.

[39]. Yalkowsky SH, Rubino JT. Solubilization by cosolvents I: Organic solutes in propylene glycol-water mixtures. J Pharm Sci, 74: 416-421, 1985.

[40]. Li A, Yalkowsky SH. Solubility of organic solutes in ethanol-water mixtures. J Pharm Sci, 83: 1735-1740, 1994.

[41]. Gould PL, Goodman M, Hanson PA. Investigation of the solubility relationships of polar, semi-polar and non-polar drugs in mixed co-solvent systems. Int J Pharm, 19: 149-159, 1984.

[42]. Chien YW, Lambert HJ. Solubilization of steroids by multiple co-solvent systems. Chem Pharm Bull, 23: 1085-1090, 1975.

[43]. Rubino JT, Blanchard J, Yalkowsky SH. Solubilization by cosolvents II: Phenytoin in binary and ternary solvents. J Parent Sci Tech, 38: 215-221, 1984.

[44]. Amidon GL, Williams NA. A solubility equation for nonelectrolytes in water. Int J Pharm, 11: 249-256, 1982.

[45]. Acree Jr WE, Rytting JH. Solubility in binary solvent systems: III. Predictive expressions based on molecular surface area. J Pharm Sci, 72: 292-296, 1983.

[46]. Bustamante P, Hinkley DV, Martin A., Shi S. Statistical analysis of the extended Hansen method using bootstrap technique. J Pharm Sci, 80: 971-977, 1991.

[47]. Reillo A, Bustamante P, Escadra B, Jimanez MM, Selles E. Solubility parameter based methods for predicting the solubility of sulfapryidine in solvent mixtures. Drug Dev Ind Pharm, 21: 2073-2084, 1995.

[48]. Jouyban-Gharamaleki A, Romero S, Bustamante P, Clark BJ. Multiple solubility maxima of oxolinic acid in mixed solvents and a new extension of Hildebrand solubility approach. Chem Pharm Bull, 48: 175-178, 2000.

[49]. Williams NA, Amidon GL. Excess free energy approach to the estimation of solubility in mixed solvent system. III. Ethanol-propylene glycol-water mixtures. J Pharm Sci, 73: 18-23, 1984.

[50]. Barzegar-Jalali M, Jouyban-Gharamaleki A. Models for calculating solubility in binary solvent systems. Int J Pharm, 140: 237-248, 1996

[51]. Jouyban A, Romero S, Chan HK, Clark BJ, Bustamante P. A cosolvency model to predict solubility of drugs at several temperatures from a limited number of solubility measurements. Chem Pharm Bull, 50: 594-599, 2002.

[52]. Jouyban A, Chan HK, Clark BJ, Acree Jr WE. Mathematical representation of apparent acid dissociation constants in aqueous-organic solvent mixtures. Int J Pharm, 246: 135-142. 2002.

[53]. Jouyban-Gharamaleki A, Hanaee J. A. novel method for improvement of predictability of the CNIBS/R-K equation. Int J Pharm, 154: 245-247, 1997.

[54]. Zvaigzne AI, Acree Jr WE. Solubility of anthracene in binary alcohol + 2-methyl-1-propanol and alcohol + 3-methyl-1-butanol solvent mixtures. J Chem Eng Data, 40: 917-919, 1995.

[55]. Acree Jr WE. Comments concerning model for solubility estimation in mixed solvent systems. Int J Pharm, 127: 27-30, 1996.

[56]. Jouyban-Gharamaleki A, Valaee L, Barzegar-Jalali M, Clark BJ, Acree Jr WE. Comparison of various cosolvency models for calculating solute solubility in water-cosolvent mixtures. Int J Pharm, 177: 93-101, 1999.

[57]. Jouyban-Gharamaleki A, Acree Jr WE. Comparison of models for describing multiple peaks in solubility profiles. Int J Pharm, 167: 177-182, 1998.

[58]. Jouyban-Gharamaleki A, Khaledi MG, Clark BJ. Calculation of electrophoretic mobilities in water-organic modifier mixtures in capillary electrophoresis. J Chromatogr A, 868: 277-284, 2000.

[59]. Jouyban A, Grosse SC, Chan HK, Coleman MW, Clark BJ. Mathematical representation of electrophoretic mobility of basic drugs in ternary solvent buffers in capillary zone electrophoresis. J Chromatogr A, 994: 191-198, 2003.

[60]. Jouyban A, Khoubnasabjafari M, Chan HK, Altria KD, Clark BJ. Predicting electrophoretic mobility of beta-blockers in water-methanol mixed electrolyte system. Chromatographia, 57: 191-196, 2003.

[61]. Jouyban A, Chan HK, Barzegar-Jalali M, Acree Jr WE. A model to represent solvent effects on the chemical stability of solutes in mixed solvent systems. Int J Pharm, 243: 167-172, 2002.

[62]. Jouyban A, Soltani S, Chan HK, Acree Jr WE. Modeling acid dissociation constant of analytes in binary solvents at various temperatures using Jouyban-Acree model. Thermochim Acta, 428: 119-123, 2005.

[63]. Jouyban A, Rashidi MR, Vaez-Gharamaleki Z, Matin AA, Djozan Dj. Mathematical representation of analyte's capacity factor in binary solvent mobile phases using Jouyban-Acree model. Pharmazie, 60: 827-829, 2005.

[64]. Jouyban A, Fathi-Azarbayjani A, Barzegar-Jalali M, Acree Jr WE. Correlation of surface tension of mixed solvents with solvent composition. Pharmazie, 59: 937-941, 2004.

[65]. Jouyban A, Khoubnasabjafari M, Vaez-Gharamaleki Z, Fekari Z, Acree Jr WE. Calculation of the viscosity of binary liquids at various temperatures using Jouyban-Acree model. Chem Pharm Bull, 53: 519-523, 2005.

[66]. Jouyban A, Fathi-Azarbayjani A, Khoubnasabjafari M, Acree Jr. WE. Mathematical representation of the density of liquid mixtures at various temperatures using Jouyban-Acree model. Indian J Chem A, 44: 1553-1560, 2005.

[67]. Jouyban A, Khoubnasabjafari M, Acree Jr WE. Modeling the solvatochromic parameter (E_T^N) of mixed solvents with respect to solvent composition and temperature using Jouyban-Acree model. Daru, 14: 22-25, 2006.

[68]. Jouyban A, Soltani S, Khoubnasabjafari M, Acree Jr WE. Refractive index correlation of solvent mixtures at various temperatures. Asian J Chem, 18: 2037-2040, 2006.

[69]. Hasan M, Shirude DF, Hiray AP, Sawant AB, Kadam UB. Densities, viscosities and ultrasonic velocities of binary mixtures of methylbenzene with hexan-2-ol, heptan-2-ol and octan-2-ol at T = 298.15 and 308.15 K. Fluid Phase Equilib, 252: 88-95, 2006.

[70]. Faller B, Ertl P. Computational approaches to determine drug solubility. Adv Drug Deliv Rev, 59: 533-545, 2007.

[71]. Sepassi K, Yalkowsky SH. Solubility prediction in octanol: A technical note. AAPS Pharm Sci Technol, 7: article 26, 2006.

[72]. Delaney JS. Predicting aqueous solubility from structure. Drug Discov Today, 10: 289-295, 2005.

[73]. Strickley RG. Solubilizing excipients in oral and injectable formulations. Pharm Res, 21: 201-230, 2004.

[74]. Sweetana S, Akers MJ. Solubility principles and practices for parentral drug dosage form development. PDA J Pharm Sci Technol, 50: 330-342, 1996.

[75]. Jouyban-Gharamaleki A, York P, Hanna M, Clark BJ. Solubility prediction of salmeterol xinafoate in water-dioxane mixtures. Int J Pharm, 216: 33-41, 2001.

[76]. Jouyban A. In silico prediction of drug solubility in water-dioxane mixtures using Jouyban-Acree model. Pharmazie, 62: 46-50, 2007.

[77]. Jouyban A. Solubility prediction of drugs in water-PEG 400 mixtures. Chem Pharm Bull, 54: 1561-1566, 2006

[78]. Jouyban A. Prediction of drug solubility in water-propylene glycol mixtures using Jouyban-Acree model. Pharmazie, 62: 365-367, 2007.

[79]. Jouyban A, Fakhree MAA, Hamzeh-Mivehroud M, Acree Jr WE. Modelling the deviations of solubilities in water-dioxane mixtures from predicted solubilities by the Jouyban-Acree model. J Drug Deliv Sci Technol, 17: 359-363, 2007

[80]. Jouyban A, Fakhree MAA, Ghafourian T, Saei AA, Acree Jr WE. Deviations of drug solubility in water-cosolvent mixtures from the Jouyban-Acree model. Effect of solute structure. Pharmazie, 63: 113-121, 2008.

[81]. Jouyban A, Soltanpour Sh, Soltani S, Acree Jr WE. Solubility prediction of drugs in water-cosolvent mixtures using Abraham solvation parameters. J. Pharm Pharmaceut Sci, 10: 263-277, 2007.

[82]. Jouyban-Gharamaleki A, Clark BJ, Acree Jr WE. Prediction of drug solubility in ternary solvent mixture. Drug Dev Ind Pharm, 26: 971-973, 2000.

[83]. Jouyban A, Chan HK, Chew NYK, Khoubnasabjafari M, Acree Jr WE. Solubility prediction of paracetamol in binary and ternary solvent mixtures using Jouyban-Acree model. Chem Pharm Bull, 54: 428-431, 2006.

[84]. Jouyban A, Chew NYK, Chan HK, Khoubnasabjafari M, Acree Jr WE. Solubility predicton of salicylic acid in water-ethanol-propylene glycol mixtures using the Jouyban-Acree model. Pharmazie, 61: 417-419, 2006.

[85]. Jouyban A, Fakhree MAA. A new definition of solubilization power of a cosolvent. Pharmazie, in press, 2008.

[86]. Comor JJ, Kopecni MM. Prediction of gas chromatography solute activity coefficients in mixed stationary phases based on the Wilson equation. Anal Chem, 62: 991-994, 1990.

[87]. Wu PL, Martin A. Extended-Hildebrand solubility approach: p-hydroxybenzoic acid in mixtures of dioxane and water. J Pharm Sci, 72: 587-592, 1983.

[88]. Tarantino R, Bishop E, Chen F-C, Iqbal K, Malick W. N-Methyl-2-pyrrolidone as a cosolvent: Relationship of cosolvent effect with solute polarity and the presence of proton-donating groups on model drug compounds. J Pharm Sci, 83: 1213-1216, 1994.

[89]. Rytting E, Lentz KA, Chen XQ, Qian F, Venkatesh S. A qantitative structure-property relationship for predicting drug solubility in PEG 400/water cosolvent systems. Pharm Res, 21: 237-244, 2004.

[90]. Jouyban A, Majidi MR, Jalilzadeh H, Asadpour-Zeynali K. Modeling drug solubility in water-cosolvent mixtures using an artificial neural network (ANN). Farmaco 59: 505-512, 2004.

[91]. Jouyban A, Chew NYK, Chan HK, Sabour M, Acree Jr WE. A unified cosolvency model for calculating solute solubility in mixed solvents. Chem Pharm Bull, 53: 634-637, 2005.

[92]. Ikeda H, Chiba K, Kanou A, Hirayama N. Prediction of solubility of drug by conductor-like screening model for real solvents. Chem Pharm Bull, 53: 253-255, 2005.

[93]. Machatha SG, Bustamante P, Yalkowsky SH. Deviation from linearity of drug solubility in ethanol/water mixtures. Int J Pharm, 283: 83-88, 2004.

[94]. Machatha SG, Yalkowsky SH. Bilinear model for the prediction of drug solubility in ethanol/water mixtures. J Pharm Sci, 94: 2731-2734, 2005.

[95]. Machatha SG, Yalkowsky SH. Estimation of the ethanol/water solubility profile from the octanol/water partition coefficient. Int J Pharm, 286: 111-115, 2004.

[96]. Myrdal PB, Manka AM, Yalkowsky SH. AQUAFAC: Aqueous functional group activity coefficients: Application to the estimation of aqurous solubility. Chemosphere, 30: 1619-1637, 1995.

[97]. Lee YC, Myrdal PB, Yalkowsky SH. AQUAFAC: Aqueous functional group activity

coefficients: Application to complex organic compounds. Chemosphere, 33: 2129-2144, 1998.

[98]. Jouyban-Gharamaleki A. PhD Dissertation, University of Bradford, UK, 2001.

[99]. Yalkowsky SH, Valvani SC. Solubility and partitioning I: Solubility of nonelectrolytes in water. J Pharm Sci, 68: 912-922, 1980.

[100]. Huuskonen J, Rantanen J, Livingstone D. Prediction of aqueous solubility for a diverse set of organic compounds based on atom-type electrotopological indices. Eur J Med Chem, 35: 1081-1088, 2000.

[101]. Abraham MH, Le J. The correlation and prediction of the solubility in water using an amended solvation energy relationship. J Pharm Sci, 88: 868-880, 1999.

[102]. Jain N, Yalkowsky SH. Estimation of the aqueous solubility I: Applications to organic nonelectrolytes. J Pharm Sci, 79: 234-252, 2001.

[103]. Dickhut RM, Armstrong DE, Andren AW. The solubility of hydrophobic aromatic chemicals in organic solvent/water mixtures: Evaluation of four mixed solvent solubility estimation methods. Environ Tox Chem, 10: 881-889, 1991.

[104]. Bodor N, Huang MJ. A new method for the estimation of the aqueous solubility of organic compounds. J Pharm Sci, 81: 954-960, 1992.

[105]. Sutter JM, Jurs PC. Prediction of aqueous solubility for a diverse set of heteroatom-containing compounds using quantitative structure property relationship. J Chem Inf Comput Sci, 36: 100-107, 1996.

[106]. Pinsuwan S, Myrdal PB, Lee YC, Yalkowsky SH. AQUAFAC 5: Aqueous functional group activity coefficients: Applications to alcohols and acids. Chemosphere, 35: 2503-2513, 1997.

[107]. Reulle P, Kesselring UW. Aqueous solubility prediction of environmental important chemicals from the mobile order thermodynamics. Chemosphere, 34: 275-298, 1997.

[108]. Subrahmanyam CVS, Srenivasa M, Venkata Rao J, Gundu Rao P. Irregular solution behaviour of paracetamol in binary solvents. Int J Pharm, 78: 17-24, 1992.

[109]. Regosz A, Pelplinska T, Kowalski P, Thiel Z. Prediction of solubility of sulfonamides in water and organic solvents based on the extended regular solution theory. Int J Pharm, 88: 437-442, 1992.

[110]. Bustamante P, Romero S, Reillo A. Thermodynamics of paracetamol in amphiprotic and amphiprotic-aprotic solvent mixtures. Pharm Sci, 1: 505-507, 1995.

[111]. Dwivedi SK, Sattari S, Jamali F, Mitchell A.G. Ibuprofen racemate and enantiomers: Phase diagram, solubility and thermodynamic studies. Int J Pharm, 87: 95-104, 1992.

[112]. Kishi H, Hashimoto Y. Evaluation of the procedure for the measurement of water solubility and n-octanol/water partition coefficient of chemicals. Results of a ring test in Japan. Chemosphere, 18: 1749-1759, 1989.

[113]. Bustamante P, Romero S, Pena A, Escalera B, Reillo A. Enthalpy-entropy compensations for the solubility of drugs in solvent mixturs; paracetamol, acetanilide and nalidixic acid in dioxane-water. J Pharm Sci, 87: 1590-1596, 1998

[114]. Kulkarni AR, Soppimath KS, Dave AM, Mehta MH, Aminabhavi TM. Solubility study of hazardous pesticide (chlorpyrifos) by gas chromatography. J Hazardous Materials A, 80: 9-13, 2000.

[115]. Roy D, Ducher F, Laumain A, Legendre JY. Determination of the aqueous solubility of drugs using a convenient 96-well plate based assay. Drug Dev Ind Pharm, 27: 107-109, 2001.

[116]. Romero S, Escalera B, Bustamante P. Solubility behaviour of polymorphs I and II of mefenamic acid in solvent mixtures. Int J Pharm, 178: 193-202, 1999.

[117]. Jouyban-Gharamaleki A, Dastmalchi S, Chan HK, Hanaee J, Javanmard A, Barzegar-Jalali M. Solubility prediction for furosemide in water-cosolvent mixtures using the minimum number of experiments. Drug Dev Ind Pharm, 27: 577-583, 2001.

[118]. Jouyban A, Chan HK, Romero S, Khoubnasabjafari M, Bustamante P. Solubility prediction in water-ethanol mixtures based on the excess free energy approach using a minimum number of experimental data. Pharmazie, 59: 117-120, 2004.

[119]. Ruckenstein E, Shulgin I. Solubility of drugs in aqueous solutions. 2. Binary nonideal mixed solvents. Int J Pharm, 260: 283-291, 2003.

[120]. Mirmehrabi M, Rohani S, Perry L. Thermodynamic modeling of activity coefficient and prediction of solubility: Part 2. Semipredictive or semiempirical models. J Pharm Sci, 95: 798-809, 2006.