Refinement of the theoretical solubility model and prediction of solute solubility in mixed solvent systems

Yueqiang Zhao [a], *, Weiwei Liu [a], Xiaoqin Pei [a], Dongrui Yao [b]

[a] Department of Chemical Engineering, Co-Innovation Center of Jiangsu Marine Bio-industry Technology, Huaihai Institute of Technology, Lianyungang 222005, Jiangsu, People's Republic of China
[b] Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing 210014, China

ABSTRACT

We have refined the non-polar (van der Waalls) and hydrogen bonding interaction term of STMIPE-AC (statistical thermodynamic molar intermolecular potential energy activity coefficient) model, which reduces the average error of the model in correlating the molar internal energy change on vaporization of pure liquids (62 solvents, 10 temperature points for each solvent) from 6.6% to 3.1%. The binary interaction parameters k_{ij} are fitted to the solubility data of solid solutes in pure solvents or estimated from a semi-empirical equation using a similar local composition method, and then they are used to predict the solute solubility in mixed solvents. The overall average relative error (ARE) in predicted solubility (15 solutes, 64 systems, 1091 data points) is 11.24%, which makes this model a practical approach for crystallization process development.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

At the early stage of a drug process development, the solubility data is usually used to select appropriate solvents for the purification of crude product of pharmaceuticals [1—3]. The large number of solvents makes it a very time and product consuming task to determine the solubility of drug in different solvents experimentally. When the solubility data is unavailable; prediction of drug solubility in various solvents is usually performed by Solid-Liquid Equilibrium equation based on activity coefficient models, such as UNIFAC [4,5], COSMO-RS [6,7], COSMO-SAC [8—11], NRTL-SAC [12,13] and STMIPE-AC (statistical thermodynamic molar intermolecular potential energy activity coefficient) model [14,15]. Since drugs are usually complex and flexible molecules containing many chemical groups, this makes it difficult to quantify interactions for drug molecules, especially for hydrogen bonding interactions; moreover, some proton acceptor sites of hydrogen bonding are not accessible due to steric hindrance of molecules; all of this will result in large errors when predicting the activity coefficient of solid solutes in solvents by thermodynamic models. In recent years, the Statistical Associating Fluid Theory (SAFT) [16—18], Cubic Plus Association equation of state (CPA) [19,20], nonrandom hydrogen-bonding (NRHB) equation of state [21], Hansen Solubility Parameters (HSP) method [22,23] and Flory—Huggins model based on the Hansen Solubility Parameters (FH-HSP) [24] have made great success in calculation of thermodynamic properties for associating (hydrogen-bonding) fluids.

Our previous work [15] presented a new theoretical solubility model, which uses the statistical mechanical expression of activity coefficient and an estimation equation of molar intermolecular potential energy for species in liquid mixtures. The model equation of molar intermolecular potential energy consists of van der Waals (hard core repulsion and dispersive cohesion), electrostatic (dipole-dipole, dipole-quadrupole and quadrupole-quadrupole interactions) and hydrogen bonding interaction terms. This activity coefficient model satisfying the thermodynamic consistency test is reliable for VLE calculations, the overall average relative error (ARE) of this theoretical model in predicting the solubility of solutes in pure organic solvents is 17.4%. In this work, we first refine the model equation of molar intermolecular potential energy by using the non-polar surface area instead of total surface area of molecule in computing the van der Waals interaction energy term, and introducing a geometric damping factor f_{geo} describing the steric hindrance of proton acceptor sites for hydrogen bonding interactions; then extend the theoretical solubility model from binary mixtures to multi-component mixtures, and verify the applicability

of this improved solubility model for practical uses.

2. Theory

2.1. Solubility predictions

The solubility of a solid solute (component 1) in mixed solvents is determined by the equation of Solid-Liquid Equilibrium (SLE) [1–3].

$$\ln(\gamma_1 x_1) = -\frac{\Delta H_m}{RT}\left(1 - \frac{T}{T_m}\right) \tag{1}$$

where x_1 is mole fraction solubility of the solute; ΔH_m and T_m are the molar melting enthalpy and melting point of solid solute correspondingly; R is the gas constant; T is the temperature; γ_1 is the activity coefficient of solute calculated by the statistical mechanical expression [14,15].

$$\begin{aligned}
\gamma_i &= \frac{V_i^{(o)}}{V^{(s)}}\exp\left[\frac{\left(<u_{inter,i}^{(s)}> - <u_{inter,i}^{(o)}>\right)}{kT}\right] \\
&= \frac{V_i^{(o)}}{V^{(s)}}\exp\left[\frac{\left(U_{inter,i}^{(s)} - U_{inter,i}^{(o)}\right)}{RT}\right]
\end{aligned} \tag{2}$$

where k is Boltzmann's constant, $V^{(s)}$ is the molar volume of liquid mixture (solution), $V_i^{(o)}$ is the molar volume of pure liquid i at system temperature and pressure, the superscripts "s"and "o" stand for solution (mixture) and pure fluid correspondingly, $<u_{inter,i}^{(s)}>$ is the ensemble average intermolecular potential energy experienced by one molecule of species i in liquid mixture, $<u_{inter,i}^{(o)}>$ is the mean intermolecular potential energy for one molecule of pure liquid i at system temperature and pressure. $U_{inter,i}^{(s)}$ is the molar intermolecular potential energy of component i in liquid mixture, $U_{inter,i}^{(s)} = N_A <u_{inter,i}^{(s)}>$, N_A is Avogadro's constant; $U_{inter,i}^{(o)}$ is the molar intermolecular potential energy of pure liquid i at system temperature and pressure, $U_{inter,i}^{(o)} = N_A <u_{inter,i}^{(o)}> = -E_i$, E_i is the molar internal energy change on vaporization (or total cohesion energy) of pure liquid i defined as [3].

$$E_i = \Delta H_{vap,i}^{(o)} - RT \tag{3a}$$

where $\Delta H_{vap,i}^{(o)}$ is the molar vaporization enthalpy of pure liquid i at system temperature, which is derived from the Clausius-Clapeyron Equation [3].

$$\Delta H_{vap,i}^{(o)} = TV_i^{(V,o)}\frac{dp_i^s}{dT} \tag{3b}$$

where $V_i^{(o,V)}$ is the vapor phase molar volume of pure species i; p_i^s is the saturated-vapor pressure of pure liquid i at system temperature, which can be calculated by Vapor-Pressure equation (i.e., Antoine equation) [25].

We usually compute the volume ratio of pure liquid to mixture approximately

$$\frac{V_i^{(o)}(T)}{V^{(s)}(T)} \approx \frac{V_i^{(o)}(T_0)}{V^{(s)}(T_0)} \tag{4a}$$

$$V^{(s)}(T_0) \approx \sum_{j=1}^{nc} V_j^{(o)}(T_0)x_j \tag{4b}$$

where T_0 is the reference temperature ($T_0 = 293.15$ K), x_j is the mole fraction of component j, n_c is the number of components, $V_j^{(o)}(T_0)$ is the molar volume of pure liquid j at temperature T_0.

2.2. Prediction of the molar intermolecular potential energy of pure fluids

The hydrogen-bond interactions play an important role for the solubility behavior of drugs in polar and associating solvent mixtures. The hydrogen bond [26–28] is a D-H …:A weak link that connects a proton donor (covalently bound hydrogen atom) to a proton acceptor or lone-pair carrier (an electronegative atom or a multiple π bond). The proton acceptors [26–28] are in the upper right-hand corner of the periodic table (C, N, O, S, halogens and π-bonds). In this refinement, we propose to divide the total surface area of a molecule into two parts: non-polar surface area and polar surface area, $A_{COSMO} = A_{nonpolar} + A_{polar}$, only the non-polar surface area contributes to the van der Waalls (hard core repulsion and dispersive cohesion) interactions, while the polar surface area is concerning with the coulombic (dipole-dipole, dipole-quadrupole and quadrupole-quadrupole interactions) and hydrogen bonding interactions.

For non-electrolyte (neutral) fluids, the molar intermolecular potential energy of a pure fluid contains contributions from non-polar (van der Waalls), polar (coulombic) and hydrogen bonding interactions [15].

$$U_{inter,i}^{(o)} = U_{vdw,i}^{(o)} + U_{cou,i}^{(o)} + U_{hb,i}^{(o)} \tag{5a}$$

$$U_{vdw,i}^{(o)} = 432Rc_{vdw}\eta A_{nonpolar,i}^{0.5}\left(R_{vdw}^{12} - R_{vdw}^6\right)f(T) \tag{5b}$$

$$\begin{aligned}
U_{cou,i}^{(o)} &= U_{DD}^{(o)} + U_{DQ}^{(o)} + U_{QQ}^{(o)} \\
&= -\frac{4.7106144 \times 10^{10}\eta}{T_0}\left(\frac{\mu_i^4}{3R_{cou1}^6} + \frac{\mu_i^2 Q_i^2}{R_{cou2}^8} + \frac{7Q_i^4}{5R_{cou3}^{10}}\right)f(T)
\end{aligned} \tag{5c}$$

$$U_{hb,i}^{(o)} = -C_{hb}n_{hb}P_{acc}P_{don}\eta\frac{V_{cosmo,i}^{1/3}}{R_{hb}}\frac{(X_{acc} - X_{don})}{X_{don}}f(T)f_{geo}^2 \tag{5d}$$

where $T_r = T/T_c$, $T_{r0} = T_0/T_c$, $R_{vdw} = \frac{(\sigma - 2a)}{(\xi_{vdw}d_{av} - 2a)}$, $R_{hb} = d_{av} + a_{hb}/d_{av}$, $f(T) = \left(\frac{1-T_r}{1-T_{r0}}\right)^{n_T}$, $R_{cou1} = \xi_{cou}d_{av} + a_{c1}/d_{av}$, $R_{cou2} = \xi_{cou}d_{av} + a_{c2}/d_{av}$, $R_{cou3} = \xi_{cou}d_{av} + a_{c3}/d_{av}$, $U_{vdw,i}^{(o)}$, $U_{cou,i}^{(o)}$ and $U_{hb,i}^{(o)}$ represent the non-polar, coulombic and hydrogen bonding interactions term of molar intermolecular potential energy correspondingly; $U_{DD}^{(o)}$, $U_{DQ}^{(o)}$ and $U_{QQ}^{(o)}$ refer to molar intermolecular dipole-dipole, dipole-quadrupole and quadrupole-quadrupole interaction energy correspondingly, μ_i and Q_i are the dipole and quadrupole moment of molecule i correspondingly, A_{COSMO} and V_{COSMO} are total cavity surface area and total cavity volume of molecule correspondingly, M_w is molecular weight, V_c is the critical volume, T_c is the critical temperature, a is Kihara hard core radius, α_s is a sphericity parameter