JOURNAL OF CHEMICAL & ENGINEERING DATA

ARTICLE

pubs.acs.org/jced

A Predictive Model for the Solubility and Octanol−Water Partition Coefficient of Pharmaceuticals

Chieh-Ming Hsieh,[†] Shu Wang,[‡] Shiang-Tai Lin,[*,†] and Stanley I. Sandler[‡]

[†]Department of Chemical Engineering, National Taiwan University, Taipei, 10617, Taiwan

[‡]Center for Molecular and Engineering Thermodynamics, Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716, United States

Supporting Information

ABSTRACT: The prediction of drug solubility in various pure and mixed solvents and the octanol−water partition coefficient (K_{OW}) are evaluated using the recently revised conductor-like screening segment activity coefficient (COSMO-SAC) model. The solubility data of 51 drug compounds in 37 different solvents and their combinations over a temperature range of 273.15 K to 323.15 K (300 systems, 2918 data points) are calculated from the COSMO-SAC model and compared to experiments. The solubility data cover a wide range of solubility from (10^{-1} to 10^{-6}) in mole fraction. When only the heat of fusion and the normal melting temperature of the drug are used, the average absolute error from the revised model is found to be 236 %, a significant reduction from that (388 %) of the original COSMO-SAC model. When the pure drug properties (heat of fusion and melting temperature) are not available, predictions can still be made with a similar accuracy using the solubility data of the drug in any other solvent or solvent mixture. The accuracy in prediction of the solubility in a mixed solvent can be greatly improved (average error of 70 %), if the measured solubility data in one pure solvent is used. Also, the standard deviation in log K_{OW} of 89 drugs, whose values range from −3.7 to 5.25, is found to be 1.14 from the revised COSMO-SAC model. Our results show that the COSMO-SAC model can provide reliable predictions of properties of pharmaceuticals and is a useful tool for drug discovery.

1. INTRODUCTION

The knowledge of solubility and the octanol−water partition coefficient (K_{OW}) of a drug are important in drug discovery, development, and manufacturing.[1−3] For example, in the area of rational drug design, a strategy in the discovery of new drugs is to use information about the structure of a drug receptor or ligands to identify candidate drugs.[4] The octanol−water partition coefficient may be used for this process, as it is one index in the rule of five, in particular that log K_{OW} should not exceed five for a candidate drug. Another example of the use of the octanol−water partition coefficient is in quantitative structure−activity relationships (QSARs), where it is assumed that there is a simple relationship between K_{OW} and biological responses, such as the lethal dose for 50 % of test organisms, LD_{50}.[4] For the synthesis, production, and formulation of a drug, solubility is an important factor, as solubility requirements can be very different in different stages of the process. For example, high (or intermediate) solubility is desired in the reaction stage, while low solubility in the solvent is desired in the purification and crystallization stage. The solubility of a drug in water is also important to estimate the drug access to biological membranes.[5] Solvent screening (i.e., finding the optimal solvent or solvent combinations for desired solubility) requires making many measurements, a costly and time-consuming task. Although the techniques of solubility and K_{OW} measurement have improved,[6−8] it is impractical to measure these data for all drug candidates at all possible operating conditions (temperature, composition of a mixed solvent, etc.). Furthermore, as a result of the wide range of values of solubility and K_{OW}, especially in cases of extremely large or small values, the accuracy of the measurements is sometimes questionable. It is not uncommon that reported values for the same compound may differ by a factor of 10 (one log unit) or more. Thus, a predictive thermodynamic model, that does not rely on experimental data, would be very useful for the design and development of drugs.[9]

In the literature several ways to estimate values of K_{OW}[10] and the solubility[9,11−16] of organic compounds have been reported with varying degrees of accuracy. In particular, the diverse chemical structures of drug molecules and the complexity in drug−solvent interactions present great challenges. Most predictive methods for K_{OW} are based on an arbitrarily defined set of fragments and a correlation using approximately 1000 compounds (training set) for which there were reliable experimental K_{OW} data.[17] While such structure-based methods can be very accurate, their accuracy deteriorates for compounds that are outside the training set. This is especially so for compounds having complex structures, such as pharmaceuticals, for which it may also be difficult to divide the molecule systematically into the fragments or groups that were defined in the training set. Furthermore, some methods have not been verified for calculating the solubility of drugs.

Theoretically based approaches allow for the prediction of both K_{OW} and solubility of drugs simultaneously. One very

Special Issue: John M. Prausnitz Festschrift

Received: August 27, 2010
Accepted: October 21, 2010
Published: November 11, 2010

successful example is the nonrandom two-liquid segment activity coefficient (NRTL-SAC) model of Chen and Song.[13] In this model, each compound is characterized using four species-specific parameters that can be determined from the regression of few experimental vapor−liquid equilibrium or solubility data. Once the parameters are available for both the drug and the solvent, satisfactory predictions can be achieved. This method has been shown to be a practical tool in the design of drug purification processes.[18] Another example is the conductor-like screening (COSMO)-based activity coefficient models,[11,12,16] in which the drug−solvent interactions are determined from quantum mechanical solvation calculations. Therefore, predictions can be made prior to, and in the absence of, any experimental measurements. Mullins et al.[11] examined the accuracy of the conductor-like screening segment activity coefficient (COSMO-SAC) model of Lin and Sandler[19] using 33 drugs in 37 solvents and solvent mixtures. Shu and Lin[16] have shown that the accuracy of COSMO-SAC in the prediction of drug solubility in mixed solvents can be greatly improved when experimental solubility data in relevant pure solvents are used. Kaemmerer et al.[20] used the COSMO-SAC model[21] for solvent or antisolvent screening and developed a selective crystallization process for chiral pharmaceuticals. Recently, the COSMO-SAC activity coefficient model has been improved to provide better descriptions of vapor−liquid and liquid−liquid equilibria.[22]

In this study, we examine the accuracy of the latest COSMO-SAC model [denoted here as COSMO-SAC(2010)] in the prediction of solubilities and octanol−water partition coefficients of drugs. A larger set of solubility data (51 drugs in 37 solvents and their mixtures) and the K_{OW} (of 89 drugs) is used, and the performance of the model on drugs of different isomeric structures is tested. No experimental data are needed for the prediction of K_{OW}. Heat of fusion and melting temperature are used for the prediction of drug solubilities. However, when the needed data for the drug (heat of fusion and melting temperature) are missing, we show that consistent predictions can be achieved in both pure and mixed solvents using a single solubility data point. When the correction method of Shu and Lin[16] is used, the COSMO-SAC(2010) model predictions of solubility data are within 70 % of the experiment.

2. THEORY

The solubility limit of a solid drug i in a solvent S (whose composition is denoted by $\underline{x}$) is determined from the equality of chemical potentials of the drug i in the solid (assuming a pure phase) and the liquid solvent at equilibrium, that is,

$$\mu_i^{\text{solid}}(T,P) = \mu_i^{\text{liquid}}(T,P,\underline{x}) \tag{1}$$

The chemical potential of the species in the liquid mixture is related to its activity coefficient as

$$\mu_i^{\text{liquid}}(T,P,\underline{x}) = \mu_i^{\text{liquid}}(T,P) + RT\ln x_i\gamma_i(T,P,\underline{x}) \tag{2}$$

and the difference in chemical potentials of the compound in the liquid and solid states can be estimated from its heat of fusion ($\Delta_{fus}H_i$) and normal melting temperature ($T_{m,i}$)[23]

$$\mu_i^{\text{solid}}(T,P) - \mu_i^{\text{liquid}}(T,P) \cong \Delta_{fus}H_i\left(1 - \frac{T}{T_{m,i}}\right) \tag{3}$$

The correction terms involving the heat capacities have been neglected in eq 3 because their contributions are much smaller compared to the contribution from the heat of fusion. Substituting eqs 2 and 3 into eq 1, we obtain an expression for calculating the solubility of drug i, mole fraction x_i of drug in the solvent, as[24]

$$\ln x_i = \frac{\Delta_{fus}H_i}{R}\left(\frac{1}{T} - \frac{1}{T_{m,i}}\right) - \ln\gamma_i(T,P,\underline{x}) \tag{4}$$

Therefore, to evaluate the solubility x_i, it is necessary to have $\Delta_{fus}H_i$ and $T_{m,i}$ of the drug and its activity coefficient γ_i in the liquid mixture. The experimental values for $\Delta_{fus}H_i$ and $T_{m,i}$ are used when available. Most of these data can be found in the DECHEMA[25,26] and/or National Institute of Standards and Technology (NIST)[27] databases. When experimental data are not available, the group contribution method, such as that of Chickos and Acree,[28,29] can be used (only for 2-(4-methylphenyl)acetic acid in this work).

Water and 1-octanol are partially miscible at ambient conditions. When mixed, an octanol-rich phase (containing approximately 0.725 mole fraction 1-octanol and 0.275 mole fraction water) coexists with another water-rich phase (which is nearly pure water) at equilibrium. The octanol−water partition coefficient of a drug is the ratio of its distribution in these two phases when a trace amount of the drug is added, that is,

$$K_{OW,i} = \lim_{C_i\to 0}\frac{C_i^O}{C_i^W} \tag{5}$$

where C_i^O and C_i^W are the molar concentrations of drug i in the octanol-rich and water-rich phases, respectively. In the limit of low solute concentrations, C_i^O can be estimated from the mole fraction of the drug and the total molar concentration of the liquid, that is, $C_i^O \approx x_i^O C^O$ and $C_i^W \approx x_i^W C^W$. Furthermore, the equilibrium criteria require the equivalence of chemical potential of the drug in both phases

$$\mu_i^O(T,P,\underline{x}^O) = \mu_i^W(T,P,\underline{x}^W) \tag{6}$$

Therefore (from eq 2), the mole fraction of the drug is related to the activity coefficient as

$$\frac{x_i^O}{x_i^W} = \frac{\gamma_i^W(T,P,\underline{x}^W)}{\gamma_i^O(T,P,\underline{x}^O)} \tag{7}$$

and the octanol−water partition coefficient can be determined from

$$\log K_{OW,i} = \log\left(\lim_{x_i\to 0}\frac{x_i^O C^O}{x_i^W C^W}\right) = \log\left(\frac{C^O\gamma_i^{W,\infty}}{C^W\gamma_i^{O,\infty}}\right) = \log\left(\frac{8.37\gamma_i^{W,\infty}}{55.5\gamma_i^{O,\infty}}\right) \tag{8}$$

where $\gamma_i^{O,\infty}$ and $\gamma_i^{W,\infty}$ are the infinite dilution activity coefficients of the drug in the octanol-rich phase and the water-rich phase, respectively. Under ambient conditions, the octanol-rich phase contains approximately 2.3 mol·L^{-1} 1-octanol and 6.07 mol·L^{-1} water.[30] Therefore, the total concentration C^O is 8.37 mol·L^{-1}. For pure water C^W = 55.5 mol·L^{-1}.

The activity coefficient γ_i needed in eqs 4 and 8 is determined from the COSMO-SAC model originally developed by Lin and Sandler [denoted as COSMO-SAC(2002)].[19] The COSMO-SAC

model is a refinement of the COSMO-RS model of Klamt and his colleagues.[31−33] In such models[19,21,31−35] the interactions between solute and solvent are considered as the interactions between the screening charges on the paired surface segments of the same size. These surface charges on the molecular cavity are determined from a two-step first principles calculation. First, the optimal conformation of the molecule in vacuum is determined by geometry optimization using quantum mechanical (QM) density functional theory (DFT). Then, a COSMO solvation calculation[32] is used to determine the total energy in the perfect conductor (solvent with infinite dielectric constant) and the ideal screening charges on the molecular surface. The σ-profile is the probability of finding a surface segment with screening charge density σ and defined as

$$p_i(\sigma) = \frac{A_i(\sigma)}{A_i} \tag{9}$$

where A_i and $A_i(\sigma)$ are the total surface areas of molecule i and the summation of the surface areas of all segments with charge density σ; $p_i(\sigma)$ is a specific property of each pure component. The σ-profile of a mixture is the summation of the σ-profile of each substance in the system weighted by its surface area (A_i) and mole fraction (x_i), that is,

$$p_S(\sigma) = \frac{\sum_i x_i A_i p_i(\sigma)}{\sum_i x_i A_i} \tag{10}$$

To have a better description of hydrogen-bonding (hb) interactions, Hsieh et al.[22] suggested dividing the σ-profile into three components: contributions from surfaces of non-hydrogen bonding atoms, from surfaces of hydroxyl (OH) group, and from surfaces of all other types of hydrogen bond donating and accepting atoms, that is,

$$p(\sigma) = p_0^{nhb}(\sigma) + p_0^{OH}(\sigma) + p_0^{OT}(\sigma) \tag{11}$$

where $p_0^{nhb}(\sigma)$ is the collection of all non-hydrogen-bonding segments; $p_0^{OH}(\sigma)$ are the segments on the hydroxyl groups; and $p_0^{OT}(\sigma)$ includes all of the other *hb* segments, that is the segments belong to oxygen, nitrogen, fluorine, or hydrogen atoms connected to N or F atoms. For interactions between atoms of hydrogen bonding donors and acceptors, Wang et al.[21] suggested refining the hydrogen bonding σ profiles by removing the less polarized surfaces using a Gaussian-type function

$$P^{HB}(\sigma) = 1 - \exp\left(\frac{-\sigma^2}{2\sigma_0^2}\right) \tag{12}$$

with $\sigma_0 = 0.7$ e·nm^{-2}.[21] The σ-profile becomes

$$p(\sigma) = p^{nhb}(\sigma) + p^{OH}(\sigma) + p^{OT}(\sigma) \tag{13}$$

with $p^{OH}(\sigma) = p_0{}^{OH}(\sigma)\cdot P^{HB}(\sigma)$, $p^{OT}(\sigma) = p_0{}^{OT}(\sigma)\cdot P^{HB}(\sigma)$, and $p^{nhb}(\sigma) = p_0{}^{nhb}(\sigma) + [p_0{}^{OH}(\sigma) + p_0{}^{OT}(\sigma)][1 - P^{HB}(\sigma)]$.

Once the σ-profile is established, the activity coefficient (Γ) for surface segment σ_m can be determined from

$$\ln \Gamma_S^t(\sigma_m^t) = -\ln\left\{\sum_s^{nhb,OH,OT}\sum_{\sigma_n} p_S^s(\sigma_n^s)\Gamma_S^s(\sigma_n^s) \times \exp\left[\frac{-\Delta W(\sigma_m^t,\sigma_n^s)}{RT}\right]\right\} \tag{14}$$

where the subscript S denotes the solution of interest (S = i for pure liquid i), the superscripts s and t can be nhb, OH, or OT, and the segment exchange energy ΔW measures the interaction energy between two surface segments of the same area a_{eff} with charge densities σ_m and σ_n

$$\Delta W(\sigma_m^t,\sigma_n^s) = c_{ES}(\sigma_m^t + \sigma_n^s)^2 - c_{hb}(\sigma_m^t,\sigma_n^s)(\sigma_m^t - \sigma_n^s)^2 \tag{15}$$

The c_{ES}, the electrostatic interaction parameter, is a temperature-dependent function

$$c_{ES} = A_{ES} + \frac{B_{ES}}{T^2} \tag{16}$$

where A_{ES} and B_{ES} are adjustable parameters and their values are taken from literature.[22] The hydrogen-bonding interaction parameter $c_{hb}(\sigma_m^t, \sigma_n^s)$ is independent of temperature and given by

$$c_{hb}(\sigma_m^t,\sigma_n^s) = \begin{cases} c_{OH-OH} \text{ if s = t = OH and } \sigma_m^t\cdot\sigma_n^s < 0 \\ c_{OT-OT} \text{ if s = t = OT and } \sigma_m^t\cdot\sigma_n^s < 0 \\ c_{OH-OT} \text{ if s = OH, t = OT, and } \sigma_m^t\cdot\sigma_n^s < 0 \\ 0 \text{ otherwise} \end{cases} \tag{17}$$

where the values of three hydrogen-bonding interaction parameters have been given earlier.[22] With the segment activity coefficient determined above, the activity coefficient of species i in mixture S is determined from

$$\ln \gamma_{i/S} = \frac{A_i}{a_{eff}}\sum_t^{nhb,OH,OT}\sum_{\sigma_m} p_i^t(\sigma_m)[\ln \Gamma_S^t(\sigma_m) - \ln \Gamma_i^t(\sigma_m)] + \ln \gamma_{i/S}^{comb} \tag{18}$$

where $\gamma_{i/S}^{comb}$, the combinatorial contribution to activity coefficient, is used to take into account the molecular size and shape differences of the species and is given by

$$\ln \gamma_{i/S}^{comb} = \ln\frac{\phi_i}{x_i} + \frac{z}{2}q_i \ln\frac{\theta_i}{\phi_i} + l_i - \frac{\phi_i}{x_i}\sum_j x_j l_j \tag{19}$$

with $\phi_i = (x_i r_i)/(\sum_j x_j r_j)$, $\theta_i = (x_i q_i)/(\sum_j x_j q_j)$, and $l_i = 5(r_i - q_i) - (r_i - 1)$, where x_i is the mole fraction of component *i*; r_i and q_i are the normalized volume and surface area parameters for component i (the standard volume and area used for normalization are 0.06669 nm^3 and 0.7953 nm^2); the summation is over all of the species in the mixture. This revised COSMO-SAC model is denoted as the COSMO-SAC(2010) model in this study. The values of all universal parameters in the COSMO-SAC(2010) model are given in ref 22 and have not been changed. All of the species specific quantities [A_i, r_i, q_i, $p_i(\sigma)$] are obtained from first principle COSMO calculations. Two changes have been made in COSMO-SAC(2010) compared to the original COSMO-SAC-(2002) model. First, the electrostatic interaction parameter c_{ES} is made to be temperature-dependent (eq 16). Second, the variation in the strength of hydrogen bonds formed by different types of donor−acceptor pairs is explicitly taken into account by using separate σ-profile (eq 13) and interaction parameters (eq 17).

Two correction methods were used in this study depending on the availability of experimental data. First, when the data of pure drug heat of fusion and temperature of melting ($\Delta_{fus}H_i$ and $T_{m,i}$) are missing, they can be replaced with any experimental solubility

Table 1. Comparison of Accuracy in Drug Solubility Prediction in Pure and Mixed Solvents by Different Methods

# of solvent species[a]	systems (data points)	COSMO-SAC (2002)	COSMO-SAC (2010)	COSMO-SAC (2010) + Cor
1	362 (473)	1.92 (580 %)	1.81 (514 %)	
2	287 (2361)	1.61 (401 %)	1.24 (245 %)	0.54 (72 %)
3	10 (80)	0.79 (120 %)	0.50 (65 %)	0.18 (19 %)
4	1 (4)	2.34 (933 %)	1.01 (175 %)	0.23 (25 %)
overall[c]	298 (2400)	1.59 (388 %)[b]	1.21 (236 %)[b]	0.53 (70 %)[b]

[a] Systems are categorized according to the number of solvent species in the system. [b] Numbers in the parentheses are the percent errors estimated from the RMSEs. [c] The solubility data for drugs in pure solvents are excluded during the calculation of the overall RMSE.

datum (x_i(expt)) of the same drug in either a pure or mixture solvent. The experimental data allow for the evaluation of $(\Delta_{fus}H_i/R)((1/T) - (1/T_{m,i}))$ from ln x_i(expt)$\gamma_i(T,P,$(expt)) (see eq 4), where the activity coefficient is calculated using the COSMO-SAC model for the reference solvent for which the solubility datum (x_i(expt)) is available. This term can then be used in eq 4 to determine the solubility in other solvents at the same temperature as

$$x_i(\text{calc}) = \frac{x_i(\text{expt})\gamma_i(T,P,\underline{x}(\text{expt}))}{\gamma_i(T,P,\underline{x}(\text{calc}))} \quad (20)$$

where x_i(calc) denotes the predicted solubility of drug i in the mixture of desired compositions $\underline{x}$(calc). Since this method combines the use of COSMO-SAC(2010) and a solubility datum in a reference solvent, it is denoted as COSMO-SAC(2010) + ref in this study.

Second, when the experimental solubility data in relevant pure solvents are available, they can be used to correct for any error in the COSMO-SAC model for drug−solvent interactions, and therefore, a significant improvement in the prediction accuracy can be achieved for solubility in the mixture these pure solvents. This correction method was recently proposed by Shu and Lin[16] and is briefly summarized in the Appendix. The accuracy of COSMO-SAC(2010) combined with this correction method is denoted as COSMO-SAC(2010) + Cor in this work.

3. COMPUTATIONAL DETAILS

The solubility of a drug in a solution at a given temperature is determined from solving eq 4 iteratively starting with an initial guess of γ_i = 1 (ideal condition). When available, the experimental melting temperature and the heat of fusion at melting temperature are used. Otherwise (only 2-(4-methylphenyl)acetic acid), they are estimated from the group contribution method proposed by Chickos and Acree.[28,29] For K_{OW} the infinite dilution activity coefficients of the drug are first determined in pure water and in the octanol-rich phase at 298.15 K. The value of log K_{OW} is then obtained from eq 8.

The needed σ-profiles for solvent molecules are taken from the VT-2005[36] σ-profile database, except for ethanol and 1-butoxybutane which were not fully optimized according to the work of Shu and Lin.[16] The needed σ-profiles for drug molecules are taken from the VT-2006[11] σ-profile database if available. For drugs not found in the VT-2006 database, their σ-profiles are generated from DMol[3] implemented in *Cerius*[2] according to the procedure suggested in the work of Lin and Sandler.[19] All of the values of the universal parameters in the COSMO-SAC(2002) and COSMO-SAC(2010) models are taken from Lin and Sandler[19] and Hsieh et al.,[22] respectively. No parameter adjustment was made here; and therefore the results here represent the predictive power of these models. The procedure of calculation of activity coefficient from COSMO-SAC model has been well-documented[19] and is not reproduced here.

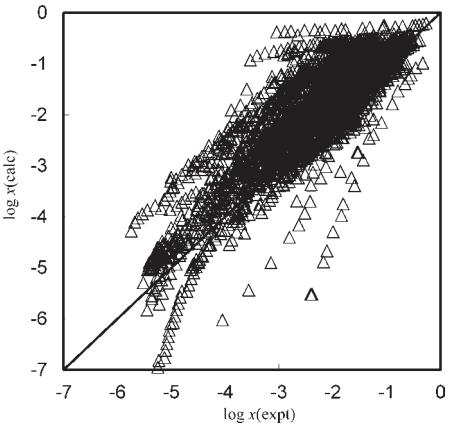

Figure 1. Comparison of drug solubility from experimental measurements[38−79] x(expt) and predictions x(calc) of the COSMO-SAC(2010) model.

4. RESULTS AND DISCUSSION

4.1. Prediction of Drug Solubility. In this study, the solubility of 52 drug compounds [from the smallest molecular iodine (2 atoms) to the largest (8*R*,9*S*,10*R*,13*S*,14*S*,17*S*)-17-hydroxy-10,13-dimethyl-1,2,6,7,8,9,11,12,14,15,16,17-dodecahydrocyclopenta[a]phenanthren-3-one, also known as testosterone (49 atoms)] in 37 different pure solvents and their mixtures are considered. There are a total of 171 drug−solvent pairs for drug solubility in pure solvent (362 systems) and 156 mixture solvent combinations (298 systems), including 152 binary solvent mixtures (287 systems), 3 ternary solvent mixtures (10 systems), and 1 quaternary solvent mixture (1 system). The temperature ranges from (273.15 to 323.15) K, and the solubility data (in mole fractions) range 10^{-1} to 10^{-6}. The complete list of the data is given in the Supporting Information. Because of the large variations in the solubility, the predicted errors are measured using the root-mean-square error (RMSE) of natural logarithm of solubility as follows[11]

$$\text{RMSE} = \left[\frac{1}{N}\sum_{i=1}^{N}(\ln x_i(\text{calc}) - \ln x_i(\text{expt}))^2\right]^{1/2} \quad (21)$$

where N is the number of data points per system [one drug in a (mixture or pure) solvent at a certain temperature]; the notations calc and expt denote the data from calculations or experiments, respectively.

dx.doi.org/10.1021/je1008872 |*J. Chem. Eng. Data* 2011, 56, 936–945

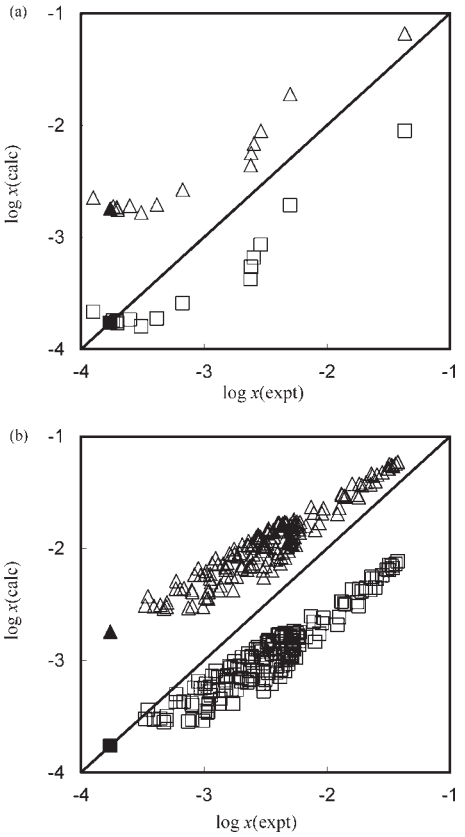

Figure 2. Comparison of the solubility of 9*H*-carbazole in (a) 15 pure solvents and (b) 19 mixed solvents from experimental measurements[69−72,78] (at 298.15 K) and predictions of COSMO-SAC(2010) [△] and COSMO-SAC(2010) + ref [□]. The reference solubility datum, marked in black symbols, is the solubility of 9*H*-carbazole in *n*-hexane at 298.15 K.

As summarized in Table 1, the overall RMSEs from COSMO-SAC(2010) for drug solubility in pure solvents is 1.81 (corresponding to 514 % in percentage error), a slight reduction from the COSMO-SAC(2002) model (1.92 or 580 %). The improvements in mixture solvent systems are more significant. The overall RMSE of drug solubility predictions in binary solvents from COSMO-SAC(2010) is 1.24 (245 %), which is only about 60 % of that from COSMO-SAC(2002) (1.61 or 401 %). Similar improvements are observed for other multisolvent systems. The RMSE of the COSMO-SAC(2002) model is similar to that in Shu and Lin's work[16] [which is 1.61 based on fewer data points (1955 data points (235 systems) for 33 drugs in 37 solvents)]. As stated in the work of Mullins et al.,[11] the COSMO-SAC(2002) model generally overpredicts the solubility. A similar tendency is observed in the predictions from the COSMO-SAC(2010) model, as shown in Figure 1. Nonetheless, the COSMO-SAC(2010) method provides a better accuracy than the COSMO-SAC(2002) model.

4.2. Prediction of Drug Solubility Using a Reference Solvent. When an experimental solubility datum of a drug in a reference (pure or mixture) solvent is available, this data point can be used to estimate the solubility in the desired (pure or mixture) solvent, eliminating the need for melting temperature and heat of fusion data for the drug. Two examples are given in this study. Figure 2 illustrates the comparison of prediction of solubility of 9*H*-carbazole in pure and mixed solvents from

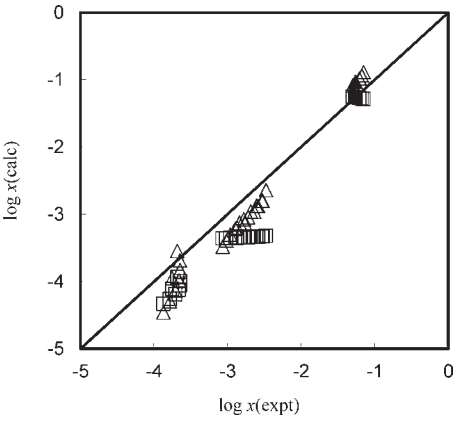

Figure 3. Comparison of solubility of paracetamol in three pure solvents from experimental measurements (temperature ranging from (273.15 to 323.15) K) and predictions of the COSMO-SAC(2010) [△] and COSMO-SAC(2010) + ref [□] models. The reference solubility datum, marked in black symbols, is the solubility of paracetamol in ethanol at 298.15 K. The data (log *x*(expt)) fall between −4 to −3.2, −3.2 to −2, and −1.5 to −0.5 corresponding to solvents of toluene, water, and ethanol, respectively.

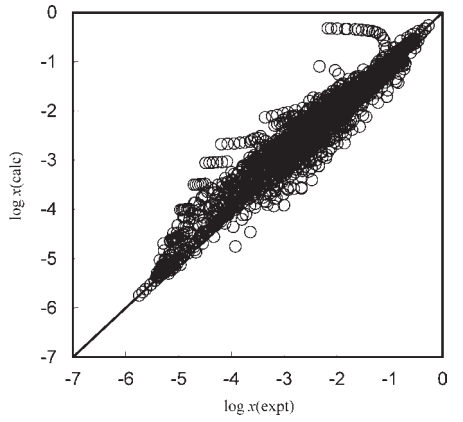

Figure 4. Comparison of drug solubility from experimental measurements[38−79] and predictions of the COSMO-SAC(2010) + Cor model.

COSMO-SAC(2010) (triangles) and COSMO-SAC(2010) + ref (squares) using the data in hexane at 298.15 K. As shown in Figure 2, the results from COSMO-SAC(2010) + ref are basically a shift of results from COSMO-SAC(2010). As described previously, the COSMO-SAC model often over predicts the solubility. The use of COSMO-SAC(2010) + ref reduces such systematic errors as seen in Figure 2.

While the correction from using a reference solvent improves the prediction of the solubility at the same temperature as the reference conditions (see eq 20), we have also tested its applicability for calculating solubilities at other temperatures. Figure 3 shows the comparison of prediction of solubility of *N*-(4-hydroxyphenyl) acetamide (a.k.a. paracetamol) in three pure solvents from COSMO-SAC(2010) and COSMO-SAC(2010) + ref in the temperature range from (273.15 to 323.15) K. It can be seen that the predicted temperature dependence of solubility varies from solvent to solvent (good agreement for toluene, lesser agreement for water and ethanol).

4.3. Improved Prediction of Drug Solubility in a Mixed Solvent by Using Solubility Data in a Pure Solvent. With the experimental drug solubility data in the relevant pure solvents,

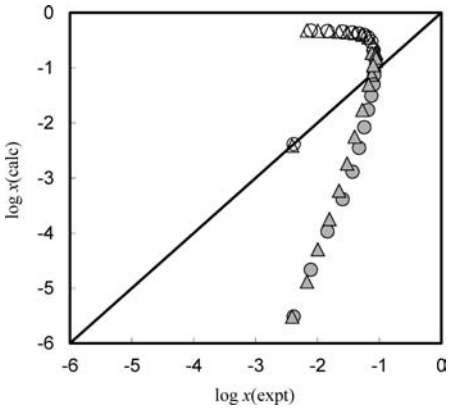

Figure 5. Comparison of solubility of 4-(dimethylamino)-1,5-dimethyl-2-phenylpyrazol-3-one (a.k.a. aminopyrine) in water + ethanol (○) and water + 1,4-dioxane (△) mixed solvents at 298.15 K. The filled and open symbols represent the results from COSMO-SAC(2010) and COSMO-SAC(2010) + Cor models. The experimental data are taken from the works of Paruta[43] and Paruta and Irani.[45]

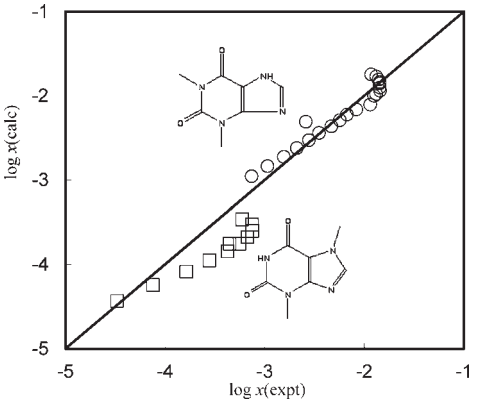

Figure 6. Comparison of solubility of 3,7-dimethylpurine-2,6-dione (□) and 1,3-dimethyl-7*H*-purine-2,6-dione (○) in water + 1,4-dioxane mixtures at 298.15 K. The experimental data are taken from the works of Martin and his colleagues.[56,57]

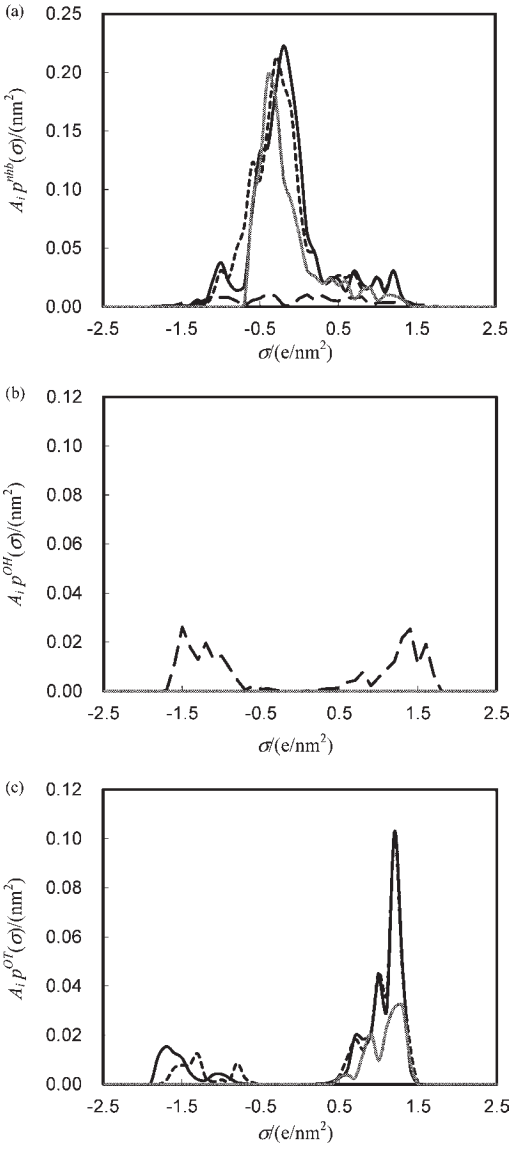

Figure 7. σ-profiles of (a) non-hydrogen-bonding (nhb), (b) hydroxyl (OH), and (c) other (OT) components for 3,7-dimethylpurine-2,6-dione (dotted lines), 1,3-dimethyl-7*H*-purine-2,6-dione (solid lines), water (long dashed lines), and 1,4-dioxane (gray lines).

the accuracy in the prediction of the drug solubility in mixed solvents can be greatly improved. This method is suggested recently by Shu and Lin[16] and denoted as "COSMO-SAC(2010) + Cor" here. Although this correction method requires additional input of experimental drug solubility data in the pure solvent, it can eliminate two important issues in the prediction of drug solubility. First, the error caused by uncertainties in the estimations or measurements of melting temperature and heat of fusion are reduced. Second, the issue of conformational flexibility becomes less significant.[16] As listed in Table 1, the overall RMSE in the prediction of drug solubility is only 0.53 (or 70 %) from COSMO-SAC(2010) + Cor, significantly less than that from COSMO-SAC(2010) without correction (1.21 or 236 %). Furthermore, the systematic error (overprediction) from the COSMO-SAC model is removed with such corrections. This can be seen from Figure 4 where the data points are evenly distributed along the diagonal line (as opposed to the higher point density on the upper right of Figure 1). It should be noted that COSMO-SAC(2010) + Cor may in some cases lead to worse predictions compared to that from COSMO-SAC(2010). Figure 5 is an example where the predictions become less accurate when corrections are applied. However, only 5 out of the total 298 systems were found to be less accurate using COSMO-SAC(2010) + Cor.

4.4. Solubility of Drug Isomers. One important merit of the COSMO-SAC model is its capability of distinguishing the properties of isomers. Here we illustrate this capability by providing the correct solubility of two drug isomers. Figure 6 illustrates the predicted solubility from the COSMO-SAC(2010) model for isomers 3,7-dimethylpurine-2,6-dione and 1,3-dimethyl-7*H*-purine-2,6-dione. These two drugs differ only by the positions of the secondary amino group and one of the tertiary amino groups. This structure difference leads to over an order of magnitude difference in their solubilities in water + 1,4-dioxane solvent mixtures at 298.15 K. As shown in Figure 6, this difference in solubility of these two drugs are correctly captured by the COSMO-SAC(2010) model. The σ-profiles of these two drugs and two solvents are shown in Figure 7. The most

dx.doi.org/10.1021/je1008872 |*J. Chem. Eng. Data* 2011, 56, 936–945

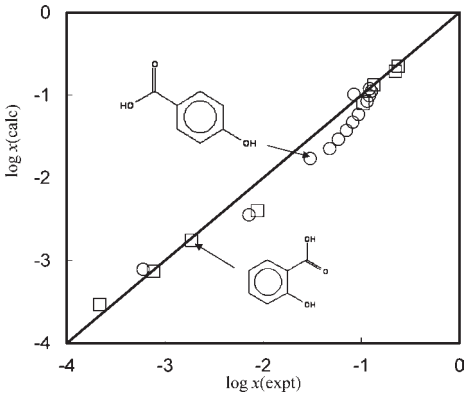

Figure 8. Comparison of the solubility of 2-hydroxybenzoic acid (□) and 4-hydroxybenzoic acid (○) in water + 1,4-dioxane solvent mixtures at 298.15 K. The experimental data are taken from the works of Wu and Martin[68] and Pena et al.[41]

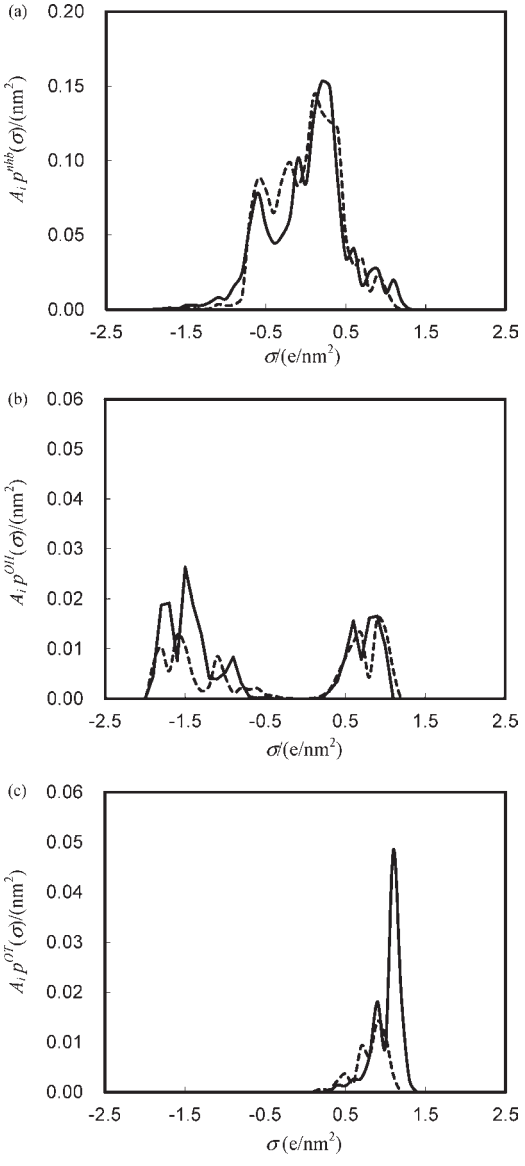

Figure 9. σ-profiles of (a) non-hydrogen-bonding (nhb), (b) hydroxyl (OH), and (c) other (OT) components for 2-hydroxybenzoic acid (dotted lines) and 4-hydroxybenzoic acid (solid lines).

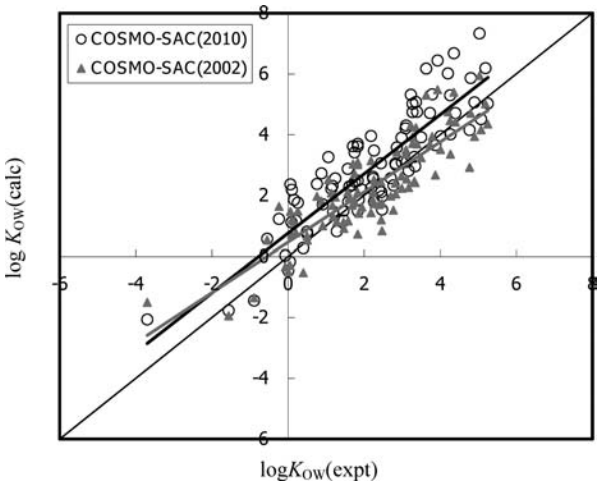

Figure 10. Comparison of measured octanol−water partition coefficients for 89 commercial drugs compared with predictions from the COSMO-SAC(2010) (○) and COSMO-SAC(2002) (△) models. The bold solid lines are the best linear fit between the predicted and the experimental data: $y = 0.98x + 0.76$ with $R^2 = 0.77$ from COSMO-SAC(2010) (dark line) and $y = 0.82\ x + 0.46$ with $R^2 = 0.75$ from COSMO-SAC(2002) (gray line).

noticeable differences in σ-profiles of these two compounds are the much more negative surface charge density on the hydrogen atom connected to nitrogen atom in 1,3-dimethyl-7*H*-purine-2,6-dione, as shown in Figure 7c. Thus, 1,3-dimethyl-7*H*-purine-2,6-dione has stronger hydrogen bonding interactions with both water and dioxane, leading to a higher solubility in both solvents and their mixtures.

Figure 8 shows the separate solubility of 2-hydroxybenzoic acid and 4-hydroxybenzoic acid in water + 1,4-dioxane solvent mixtures at 298.15 K. Different from the previous example, the differences in solubility of these two drugs in the same solvent are less significant; however, the solubility of either drug in water and in 1,4-dioxane differ by more than an order of magnitude. Therefore, the solubility varies significantly with solvent composition. Figure 9 compares the σ-profiles of these two isomers. Since the carbonyl group and hydroxyl group of 2-hydroxybenzoic acid may form an intramolecular hydrogen bond (hence less surface exposed to the solvent), their contributions to both $p^{OH}(\sigma)$ and $p^{OT}(\sigma)$ of 2-hydroxybenzoic acid are reduced compared to those of 4-hydroxybenzoic acid. Such differences result in the different solubilities observed in these two isomers.

4.5. Octanol−Water Partition Coefficient of Drug. The predictions of the logarithm of the octanol−water partition coefficients, log K_{OW}, for 89 drug compounds are shown in Figure 10. The experimental values (taken from the study of Duffy and Jorgenson[37]) for log K_{OW} of these compounds range from about −4 (2-[3,4-dihydroxy-2,5-bis(hydroxymethyl)oxolan-2-yl]oxy-6-(hydroxymethyl)oxane-3,4,5-triol, a.k.a. sucrose) to about +5 (2-[3-(trifluoromethyl)anilino]benzoic acid); that is, 10^{-4} to 10^{5} in K_{OW}. The complete list of the compounds is given in the Supporting Information. The RMSE in log K_{OW} from COSMO-SAC(2010) and COSMO-SAC(2002) are 1.14 and 0.85, respectively. While the COSMO-SAC(2010) model is found to be less accurate here, it is indeed more precise. Also shown in Figure 10 are the linear regressions of the predicted values and experimental data, that is, log K_{OW}(calc) = a logK_{OW}(expt) + b. The

dx.doi.org/10.1021/je1008872 |J. Chem. Eng. Data 2011, 56, 936–945

COSMO-SAC(2010) model has a much better linear correlation with the experimental values (larger values in the slope, 0.98 vs 0.82, and in the correlation coefficient R^2, 0.77 vs 0.75). Its lower overall accuracy is a result of the larger systematic error (seen in the intercept, 0.76 vs 0.46). The reason for the systematic errors seen in the prediction of log K_{OW} is not clear and will be the subject for further studies.

5. CONCLUSION

The COSMO-SAC activity coefficient model revised by Hsieh et al.[22] [denoted as COSMO-SAC(2010)] has been proven to provide a better description of vapor−liquid and liquid−liquid equilibria when compared with that of the original COSMO-SAC model [denoted as COSMO-SAC(2002)]. Here we show that the overall RMSE (root-mean-square error) of predictions of the COSMO-SAC(2010) model in the natural logarithm of drug solubility for 51 drugs in 37 organic solvents and their combinations is 1.21 (corresponding to 236 % in percentage error), a reduction of 25 % when compared with 1.59 (388 %) from the COSMO-SAC(2002) model. The RMSE in log K_{OW} is found to be 1.14 from COSMO-SAC(2010), slightly worse than that from COSMO-SAC(2002), 0.85. Since the parameters of COSMO-SAC were obtained without using any of the data considered in this study, the results presented here are truly a priori predictions. We also show that the RMSE of the COSMO-SAC(2010) model in the prediction of drug solubility can be significantly decreased to 0.53 (70 %) when the experimental drug solubility data in the relevant pure solvents are used. The ability of COSMO-SAC model to distinguish between isomers is also demonstrated in this study. Consequently, we believe that the COSMO-SAC-(2010) model could be a useful tool in drug discovery, purification, and formulation.

APPENDIX. CORRECTION METHOD FOR THE COSMO-SAC MODEL TO IMPROVE THE DRUG SOLUBILITY PREDICTIONS IN MIXED SOLVENT SYSTEMS WITH THE EXPERIMENTAL DRUG SOLUBILITY DATA IN THE RELEVANT PURE SOLVENTS

The COSMO-SAC model can improve the accuracy in the prediction of drug solubility when the solubility data in pure solvent is introduced. As stated in the work of Shu and Lin,[16] an empirical correction term is introduced by modifying the results of the COSMO-SAC model as follows

$$\ln \gamma_{i/S} = \ln \gamma_{i/S}^{COSMOSAC} + \ln \gamma_{i/S}^{Correction} \quad (A1)$$

where the expression of correction term is

$$\ln \gamma_{i/S}^{Correction} = \frac{1}{2RT} \sum_{j=1}^{C} \sum_{k=1}^{C} (B_{ij} + B_{ik}) x_j x_k \quad (A2)$$

where B_{ij} and B_{ik} are the binary interaction parameters between drug i and solvents j and k; C is the number of solvents in the mixture. In the case of pure solvent, eq A2 can be rewritten as

$$\ln \gamma_{i/S}^{Correction} = \frac{B_{ij}}{RT} x_j^2 = \frac{B_{ij}}{RT} (1 - x_i)^2 \quad (A3)$$

and the value of interaction parameter (B_{ij}) between drug i and solvent j can be determined by using the experimental drug solubility in pure solvent from

$$\frac{B_{ij}}{RT} = \frac{1}{(1 - x_i)^2} \left[\frac{\Delta H_{fus,i}}{R} \left(\frac{1}{T} - \frac{1}{T_{m,i}} \right) - \ln x_i \gamma_{i/S}^{COSMOSAC} \right] \quad (A4)$$

Since $\gamma_{i/S}^{COSMOSAC}$ is calculated from the COSMO-SAC model, this approach is denoted as COSMO-SAC + Cor in this study.

ASSOCIATED CONTENT

Supporting Information. A complete list of systems used in the solubility study and a comparison of experimental and predicted values of log K_{OW}. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: stlin@ntu.edu.tw.

Funding Sources

The financial support from Grant NSC 98-2221-E-002-087-MY3 by the National Science Council of Taiwan and computation resources from the National Center for High-Performance Computing of Taiwan are acknowledged. This research was also supported in part by the U.S. National Science Foundation under Grant GOALI-0853685.

ACKNOWLEDGMENT

The author would like to thank Professor David Shan-Hill Wong at the National Tsinghua University, Taiwan for his suggestion of using a reference solvent in the prediction of drug solubility.

REFERENCES

(1) Modarresi, H.; Conte, E.; Abildskov, J.; Gani, R.; Crafts, P. Model-based calculation of solid solubility for solvent selection - A review. *Ind. Eng. Chem. Res.* **2008**, *47*, 5234–5242.

(2) Constable, D. J. C.; Jimenez-Gonzalez, C.; Henderson, R. K. Perspective on solvent use in the pharmaceutical industry. *Org. Process Res. Dev.* **2007**, *11*, 133–137.

(3) Sangster, J. *Octanol-Water Partition Coefficients: Fundamentals and Physical Chemistry*; Wiley and Sons: New York, 1997.

(4) Jorgensen, W. L. The many roles of computation in drug discovery. *Science* **2004**, *303*, 1813–1818.

(5) Faller, B.; Ertl, P. Computational approaches to determine drug solubility. *Adv. Drug Delivery Rev.* **2007**, *59*, 533–545.

(6) Jordan, C. G. M.; Quigley, J. M.; Timoney, R. F. Synthesis, Hydrolysis Kinetics and Lipophilicity of *o*-Acyl Esters of Oxprenolol. *Int. J. Pharm.* **1992**, *84*, 175–189.

(7) Lambert, W. J. Modeling Oil-Water Partitioning and Membrane Permeation Using Reversed-Phase Chromatography. *J. Chromatogr., A* **1993**, *656*, 469–484.

(8) Leo, A. J. Hydrophobic Parameter - Measurement and Calculation. *Methods Enzymol.* **1991**, *202*, 544–591.

(9) Chen, C. C.; Crafts, P. A. Correlation and prediction of drug molecule solubility in mixed solvent systems with the Nonrandom Two-Liquid Segment Activity Coefficient (NRTL-SAC) model. *Ind. Eng. Chem. Res.* **2006**, *45*, 4816–4824.

(10) Ghose, A. K.; Crippen, G. M. Atomic Physicochemical Parameters for 3-Dimensional Structure-Directed Quantitative Structure-

Activity-Relationships 0.1. Partition-Coefficients as a Measure of Hydrophobicity. *J. Comput. Chem.* **1986**, *7*, 565–577.

(11) Mullins, E.; Liu, Y. A.; Ghaderi, A.; Fast, S. D. Sigma profile database for predicting solid solubility in pure and mixed solvent mixtures for organic pharmacological compounds with COSMO-based thermodynamic methods. *Ind. Eng. Chem. Res.* **2008**, *47*, 1707–1725.

(12) Klamt, A.; Eckert, F.; Hornig, M.; Beck, M. E.; Burger, T. Prediction of aqueous solubility of drugs and pesticides with COSMO-RS. *J. Comput. Chem.* **2002**, *23*, 275–281.

(13) Chen, C. C.; Song, Y. H. Solubility modeling with a nonrandom two-liquid segment activity coefficient model. *Ind. Eng. Chem. Res.* **2004**, *43*, 8354–8362.

(14) Chen, C. C.; Simoni, L. D.; Brennecke, J. F.; Stadtherr, M. A. Correlation and prediction of phase behavior of organic compounds in ionic liquids using the nonrandom two-liquid segment activity coefficient model. *Ind. Eng. Chem. Res.* **2008**, *47*, 7081–7093.

(15) Tung, H. H.; Tabora, J.; Variankaval, N.; Bakken, D.; Chen, C. C. Prediction of pharmaceutical solubility via NRTL-SAC and COSMO-SAC. *J. Pharm. Sci.* **2008**, *97*, 1813–1820.

(16) Shu, C. C.; Lin, S. T. Prediction of drug solubility in mixed solvent systems using the COSMO-SAC activity coefficient model. *Ind. Eng. Chem. Res.* **2010**No. DOI: 10.1021/ie100409y.

(17) Rekker, R. E. *The Hydrophobic Fragment Constant*; Elsevier: Amsterdam, 1976.

(18) Hsieh, D.; Marchut, A. J.; Wei, C. K.; Zheng, B.; Wang, S. S. Y.; Kiang, S. Model-based solvent selection during conceptual process design of a new drug manufacturing process. *Org. Process Res. Dev.* **2009**, *13*, 690–697.

(19) Lin, S. T.; Sandler, S. I. A priori phase equilibrium prediction from a segment contribution solvation model. *Ind. Eng. Chem. Res.* **2002**, *41*, 899–913.

(20) Kaemmerer, H.; Jones, M. J.; Lorenz, H.; Seidel-Morgenstern, A. Selective crystallisation of a chiral compound-forming system—Solvent screening, SLE determination and process design. *Fluid Phase Equilib.* **2010**, *296*, 192–205.

(21) Wang, S.; Sandler, S. I.; Chen, C. C. Refinement of COSMO-SAC and the applications. *Ind. Eng. Chem. Res.* **2007**, *46*, 7275–7288.

(22) Hsieh, C. M.; Sandler, S. I.; Lin, S. T. Improvements of COSMO-SAC for vapor-liquid and liquid-liquid equilibrium predictions. *Fluid Phase Equilib.* **2010**, *297*, 90–97.

(23) Sandler, S. I. *Chemical, Biochemical, and Engineering Thermodynamics*, 4th ed.; John Wiley & Sons, Inc.: New York, 2006.

(24) Prausnitz, J. M.; Lichtenthaler, R. N.; de Azevedo, E. G. *Molecular Thermodynamics of Fluid-Phase Equilibria*, 3rd ed.; Pearson Education Taiwan Ltd.: Taipei, 2004.

(25) Abildskov, J. *Solubility and Related Properties of Large Complex Chemicals. Part 2: Organic Solutes Ranging from C2 to C41*; DECHEMA: Frankfurt and Main, 2005; Vol. XV, Part 2.

(26) Marrero, J.; Abildskov, J. *Solubility and Related Properties of Large Complex Chemicals. Part 1: Organic Solutes Ranging from C4 to C40*. DECHEMA: Frankfurt and Main, 2003; Vol. XV, Part 1.

(27) *National Institute of Standards and Technology (NIST) webbook*. http://webbook.nist.gov/chemistry/ (accessed Aug 2010).

(28) Chickos, J. S.; Acree, W. E. Estimating solid-liquid phase change enthalpies and entropies. *J. Chem. Phys. Ref. Data* **1999**, *28*, 1535–1673.

(29) Chickos, J. S.; Acree, W. E. Total phase change entropies and enthalpies - An update on their estimation and applications to the estimations of amphiphillic fluorocarbon-hydrocarbon molecules. *Thermochim. Acta* **2003**, *395*, 59–113.

(30) Ropel, L.; Belveze, L. S.; Aki, S.; Stadtherr, M. A.; Brennecke, J. F. Octanol-water partition coefficients of imidazolium-based ionic liquids. *Green Chem.* **2005**, *7*, 83–90.

(31) Klamt, A. Conductor-Like Screening Model for Real Solvents - a New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99*, 2224–2235.

(32) Klamt, A.; Schuurmann, G. COSMO - a New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

(33) Klamt, A.; Jonas, V.; Burger, T.; Lohrenz, J. C. W. Refinement and parametrization of COSMO-RS. *J. Phys. Chem. A* **1998**, *102*, 5074–5085.

(34) Lin, S. T.; Chang, J.; Wang, S.; Goddard, W. A.; Sandler, S. I. Prediction of vapor pressures and enthalpies of vaporization using a COSMO solvation model. *J. Phys. Chem. A* **2004**, *108*, 7429–7439.

(35) Grensemann, H.; Gmehling, J. Performance of a conductor-like screening model for real solvents model in comparison to classical group contribution methods. *Ind. Eng. Chem. Res.* **2005**, *44*, 1610–1624.

(36) Mullins, E.; Oldland, R.; Liu, Y. A.; Wang, S.; Sandler, S. I.; Chen, C. C.; Zwolak, M.; Seavey, K. C. Sigma-profile database for using COSMO-based thermodynamic methods. *Ind. Eng. Chem. Res.* **2006**, *45*, 4389–4415.

(37) Duffy, E. M.; Jorgensen, W. L. Prediction of properties from simulations: Free energies of solvation in hexadecane, octanol, and water. *J. Am. Chem. Soc.* **2000**, *122*, 2878–2888.

(38) Phatak, P. V.; Gaikar, V. G. Solubilities of *o*- and *p*-chlorobenzoic acids and *o*- and *p*-nitroanilines in *N*,*N*-dimethylformamide plus water. *J. Chem. Eng. Data* **1996**, *41*, 1052–1054.

(39) Pacheco, D. P.; Martinez, F. Thermodynamic analysis of the solubility of naproxen in ethanol plus water cosolvent mixtures. *Phys. Chem. Liq.* **2007**, *45*, 581–595.

(40) Bustamante, P.; Navarro, J.; Romero, S.; Escalera, B. Thermodynamic origin of the solubility profile of drugs showing one or two maxima against the polarity of aqueous and nonaqueous mixtures: Niflumic acid and caffeine. *J. Pharm. Sci.* **2002**, *91*, 874–883.

(41) Pena, M. A.; Bustamante, P.; Escaler, B.; Reillo, A.; Bosque-Sendra, J. M. Solubility and phase separation of benzocaine and salicyclic acid in 1,4-dioxane-water mixtures at several temperatures. *J. Pharm. Sci.* **2004**, *36*, 571–578.

(42) Breon, T. L.; Paruta, A. N. Solubility profiles for several barbiturates in hydroalcoholic mixtures. *J. Pharm. Sci.* **1970**, *59*, 1306–1313.

(43) Paruta, A. N. Solubility profiles for antipyrine and aminopyrine in hydroalcoholic solutions. *J. Pharm. Sci.* **1967**, *56*, 1565–1569.

(44) Gantiva, M.; Yurquina, A.; Martinez, F. Solution Thermodynamics of Ketoprofen in Ethanol plus Water Cosolvent Mixtures. *J. Chem. Eng. Data* **2010**, *55*, 113–118.

(45) Paruta, A. N.; Irani, S. A. Dielectric solubility profiles in dioxane-water mixtures for several antipyretic drugs - Effect of substituents. *J. Pharm. Sci.* **1965**, *54*, 1334–1338.

(46) Manzo, R. H.; Ahumada, A. A. Effects of solvent medium on solubility 0.5. Enthalpic and entropic contributions to the free-energy changes of disubstituted benzene-derivatives in ethanol water and ethanol cyclohexane mixtures. *J. Pharm. Sci.* **1990**, *79*, 1109–1115.

(47) Dickhut, R. M.; Andren, A. W.; Armstrong, D. E. Naphthalene solubility in selected organic solvent-water mixtures. *J. Chem. Eng. Data* **1989**, *34*, 438–443.

(48) Jimenez, J. A.; Martinez, F. Temperature dependence of the solubility of acetaminophen in propylene glycol plus ethanol mixtures. *J. Solution Chem.* **2006**, *35*, 335–352.

(49) Jimenez, J. A.; Martinez, F. Thermodynamic study of the solubility of acetaminophen in propylene glycol plus water cosolvent mixtures. *J. Braz. Chem. Soc.* **2006**, *17*, 125–134.

(50) Bustamante, P.; Romero, S.; Pena, A.; Escalera, B.; Reillo, A. Enthalpy-entropy compensation for the solubility of drugs in solvent mixtures: Paracetamol, acetanilide, and nalidixic acid in dioxane-water. *J. Pharm. Sci.* **1998**, *87*, 1590–1596.

(51) Hojjati, H.; Rohani, S. Measurement and prediction of solubility of paracetamol in water-isopropanol solution. Part 1. Measurement and data analysis. *Org. Process Res. Dev.* **2006**, *10*, 1101–1109.

(52) Romero, S.; Reillo, A.; Escalera, B.; Bustamante, P. The behavior of paracetamol in mixtures of amphiprotic and amphiprotic-aprotic solvents. Relationship of solubility curves to specific and non-specific interactions. *Chem. Pharm. Bull.* **1996**, *44*, 1061–1064.

(53) Granberg, R. A.; Rasmuson, A. C. Solubility of paracetamol in binary and ternary mixtures of water plus acetone plus toluene. *J. Chem. Eng. Data* **2000**, *45*, 478–483.

(54) Judy, C. L.; Acree, W. E. Solubility in binary solvent systems. 5. Monomer and dimer models for the solubility of paratolylacetic acid in systems of non-specific interactions. *Int. J. Pharm.* **1985**, *27*, 39–44.

(55) Martin, A.; Miralles, M. J. Extended Hildebrand solubility approach - Solubility of tolbutamide, acetohexamide, and sulfisomidine in binary solvent mixtures. *J. Pharm. Sci.* **1982**, *71*, 439–442.

(56) Martin, A.; Newburger, J.; Adjei, A. Extended Hildebrand solubility approach - Solubility of 1,3-dimethyl-7*H*-purine-2,6-dione in polar binary solvents. *J. Pharm. Sci.* **1980**, *69*, 487–491.

(57) Martin, A.; Paruta, A. N.; Adjei, A. Extended Hildebrand solubility approach - Methylxanthines in mixed-solvents. *J. Pharm. Sci.* **1981**, *70*, 1115–1120.

(58) Martin, A.; Wu, P. L.; Adjei, A.; Mehdizadeh, M.; James, K. C.; Metzler, C. Extended Hildebrand solubility approach - Testosterone and testosterone propionate in binary solvents. *J. Pharm. Sci.* **1982**, *71*, 1334–1340.

(59) Bustamante, P.; Escalera, B. Enthalpy and entropy contributions to the solubility of sulphamethoxypyridazine in solvent mixtures showing 2 solubility maxima. *J. Pharm. Sci.* **1995**, *47*, 550–555.

(60) Martin, A.; Wu, P. L.; Velasquez, T. Extended Hildebrand solubility approach - Sulfonamides in binary and ternary solvents. *J. Pharm. Sci.* **1985**, *74*, 277–282.

(61) Reillo, A.; Escalera, B.; Selles, E. Prediction of sulfanilamide solubility in dioxane-water mixtures. *Pharmazie* **1993**, *48*, 904–907.

(62) Bustamante, P.; Ochoa, R.; Reillo, A.; Escalera, J. B. Chameleonic effect of sulfanilamide and sulfamethazine in solvent mixtures - Solubility curves with 2 maixam. *Chem. Pharm. Bull.* **1994**, *42*, 1129–1133.

(63) Martin, A.; Wu, P. L.; Adjei, A.; Lindstrom, R. E.; Elworthy, P. H. Extended Hildebrand solubility approach and the log linear solubility equation. *J. Pharm. Sci.* **1982**, *71*, 849–856.

(64) Subrahmanyam, C. V. S.; Reddy, M. S.; Rao, J. V.; Rao, P. G. Irregular solution behavior of paracetamol in binary solvents. *Int. J. Pharm.* **1992**, *78*, 17–24.

(65) Khossravi, D.; Connors, K. A. Solvent effects on chemical processes 1. Solubility of aromatic and heterocyclic-compounds in binary aqueous-organic solvents. *J. Pharm. Sci.* **1992**, *81*, 371–379.

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

(67) Acree, W. E. Thermochemical investigations of associated solutions 2. Calculation of iodine benzene equiilibrium equilibrium-constants from solute solubility in binary solvent mixtures. *Int. J. Pharm.* **1983**, *15*, 159–165.

(68) Wu, P. L.; Martin, A. Extended Hildebrand solubility approach - Para-Hydroxybenzoic acid in mixtures of dioxane and water. *J. Pharm. Sci.* **1983**, *72*, 587–592.

(69) Acree, W. E.; McCargar, J. W.; Zvaigzne, A. I.; Teng, I. L. Mathematical representation of thermodynamic properties - 9*H*-carbazole solubilities in binary alkane + dibutyl ether and alkane + tetrahydropyran solvent mixtures. *Phys. Chem. Liq.* **1991**, *23*, 27–35.

(70) McCargar, J. W.; Acree, W. E. Thermochemical investigations of associated solutions 3. Effect of the inert cosolvent on solute solvent association constants calculated from solubility measurements. *Phys. Chem. Liq.* **1987**, *17*, 123–138.

(71) Acree, W. E.; McCargar, J. W. Thermochemical investigations of associated solutions 5. Calculation of solute-solvent equilibrium-constants from solubility in mixtures containing 2 complexing solvents. *J. Pharm. Sci.* **1987**, *76*, 575–579.

(72) McCargar, J. W.; Acree, W. E. Solubility of carbazole in binary chloralkane + dibutyl ether solvent mixtures. *J. Solution Chem.* **1989**, *18*, 151–158.

(73) Adjei, A.; Newburger, J.; Martin, A. Extended Hildebrand approach - Solubility of caffeine in dioxane-water mixtures. *J. Pharm. Sci.* **1980**, *69*, 659–661.

(74) Rubino, J. T.; Obeng, E. K. Influence of solute structure on deviations from the log linear solubility equation in propylene glycol—water mixtures. *J. Pharm. Sci.* **1991**, *80*, 479–483.

(75) Acree, W. E.; Rytting, J. H. Solubilities in binary solvent systems 2. The importance of nonspecific interactions. *Int. J. Pharm.* **1982**, *10*, 231–238.

(76) Acree, W. E.; Bertrand, G. L. Thermochemical investigations of nealy ideal binary solvents 7. Monomer and dimer models for solubility of benzoic-acid in simple binary and ternary solvents. *J. Pharm. Sci.* **1981**, *70*, 1033–1036.

(77) Acree, W. E.; Rytting, J. H. Solubility in binary solvent systems 1. Specific verus nonspecific interations. *J. Pharm. Sci.* **1982**, *71*, 201–205.

(78) McCargar, J. W.; Acree, W. E. Thermochemical investigations of associated solutions 4. Calculation of 9H-carbazole-dibutyl ether association constants from measured solubility in binary solvent mixtures. *J. Pharm. Sci.* **1987**, *76*, 572–574.

(79) Acree, W. E.; Rytting, J. H. Solubility in binary solvent systems 3. Predictive expressions based on molecular-surface areas. *J. Pharm. Sci.* **1983**, *72*, 292–296.