

# A New Model for Predicting Solubilities and Enthalpies of Solution for a Given Solute in a Series of Solvents

Pilar Bustamante<sup>†</sup> and Russell S. Drago\*

Department of Chemistry, University of Florida, Gainesville, Florida 32611

Received: February 10, 1997; In Final Form: April 10, 1997<sup>⊗</sup>

The unified solvation model  $\Delta\chi = PS' + E_A^*E_B + C_A^*C_B + W$ , used to predict solvent dependence of dissolved solutes, is extended to enthalpies of solution and solubilities, by predicting the energetic component for forming a cavity in the solvent. The energy to create the cavity is given without adding a new solvent parameter using the term,  $QS'^2$ , where the  $S'^2$  values are the square of the previously reported USM solvent polarity parameters,  $S'$ .  $Q$  is the solute-dependent cavity formation term. The resultant unified solvation cavity model (USCM) is applied successfully to 350 enthalpies and free energies (solubilities) of solution in effect unifying the separate areas of solvation polarity probes and regular solution theory. The resulting empirical parameters for all 35 solutes studied are consistent with the intuitive concepts that cavity formation,  $QS'^2$ , and separation of solid or liquid solute molecules upon dissolution,  $W$ , are endothermic processes, while nonspecific solvation of the solute,  $PS'$ , is exothermic. This consistency, combined with the spectroscopic basis for  $S'$ , supports the USCM interpretation as opposed to a meaningless multiparameter fit of a data set. Early approaches to predicting solubilities and correlating measurements to single parameter solvent polarity scales fail when solute–solvent specific interactions exist. As in earlier USM studies, the magnitude of specific donor–acceptor contributions is predicted with the electrostatic-covalent model, for heats of solution and solubilities. Thus, in addition to an empirical correlation, USCM provides a quantitative separation of specific and nonspecific effects. Furthermore, the work provides a model that can be used to predict solubilities with confidence. Finally, the correlation produces solute parameters that provide new insights concerning solubility. For example, the magnitudes of  $Q$  for the cavity term of different solutes leads to the novel suggestion that solute polarity is more important than solute size in determining the energy of cavity formation.

## Introduction

There is an extensive literature involving the Hildebrand solubility parameter theory<sup>1,2</sup> for predicting solubilities. The solubility parameter,  $\delta$ , is defined as the square root of the energy of vaporization divided by the molar volume. Solubilities of a substance (A) in a solvent (B) are estimated by combining their solubility parameters to estimate cavity formation,  $\delta_B^2$ , solute molecule separation in the solvent,  $\delta_A^2$ , and the solute–solvent intermolecular interactions,  $\delta_A\delta_B$ . Extension of the model to enthalpies of solution was treated by the Scatchard–Hildebrand equation.<sup>2</sup> This approach is limited to so-called regular solutions, typically nonpolar solutes and solvents. The importance of solubility predictions to areas as diverse as coatings,<sup>3</sup> polymer blending,<sup>4</sup> drug delivery,<sup>5</sup> etc., has motivated an extensive effort to predict solubilities and to understand solution properties.

Hansen<sup>6</sup> extended the regular solution approach to polar solvents by adding parameters for polarity,  $\delta_P$ , dispersion,  $\delta_D$ , and hydrogen bonding,  $\delta_H$ . The model has met with limited success, failing when specific donor–acceptor interactions exist. A statistical mechanical approach to solubility and cavity formation is offered by scaled particle theory.<sup>7–10</sup> A hard-sphere model is employed which breaks cavity formation into a step for solute accommodation followed by a step involving solvent–solute interaction. The concentration of solute dissolved is expressed as a power series in the activity or fugacity of the solute analogous to a virial equation of state for gases. In an excellent review, solubilities of small, nonbasic, nonpolar solutes

are treated and problems associated with the extension of this model to more complex systems discussed.<sup>10</sup> These complications justify attempts to find empirical models to predict solubilities, with a long-range goal of understanding the empirical models with a statistical mechanical approach.

The influence of solvents on spectral shifts has developed as a separate area from solubility parameter theory leading to solvent polarity scales.<sup>11</sup>

Attempts<sup>12</sup> to find empirical solubility models have combined Hildebrand's solubility parameters to estimate the cavity term and solvent polarity parameters from spectral shifts to estimate solvent–solute interactions. More recently<sup>13</sup> correlation of an extensive amount of solubility data ( $\Delta G$  and  $\Delta H$ ) was reported using a multiparameter linear solvation energy approach. A combination of solubility parameters and  $E$  and  $C$  parameters<sup>14</sup> to predict the specific interaction has been suggested<sup>15–17</sup> to correlate enthalpies of mixing and solubilities of polar solutes in polar solvents.

In earlier reports from this laboratory,<sup>18</sup> a unified solvation model was offered to treat solvent influences on physicochemical properties. The key advantage of this approach is the separation of specific,  $\Delta\chi_S$ , and nonspecific,  $\Delta\chi_{NS}$ , solvation components of a solvent dependent physicochemical property,  $\Delta\chi$ . This separation leads to an increased understanding of the role of solvent on spectral shifts,<sup>18a–d</sup> rates of reaction (decarboxylation of an ion-paired anion, and electron-transfer reaction rates)<sup>18e</sup> and redox potentials.<sup>18f</sup> Nonspecific solvation is treated with the equation

$$\Delta\chi_{NS} = PS' + W \quad (1)$$

where  $\Delta\chi_{NS}$  is the nonspecific solvent-dependent physicochem-

<sup>†</sup> Department of Farmacia and Tecnologia Farmaceutica, University of Alcalá de Henares, Madrid 28871, Spain.

<sup>⊗</sup> Abstract published in *Advance ACS Abstracts*, May 15, 1997.

ical property expressed in energy units. A new set of solvent polarity parameters,  $S'$ , result that are devoid of the specific interactions included in most other solvent polarity scales.  $P$  measures the susceptibility of the solute to nonspecific solvation, and  $W$  is the value of  $\Delta\chi$  when  $S' = 0$ . Contributions to solvation from specific interactions are absent in  $PS'$  and are treated with the  $E$  and  $C$  approach:<sup>14</sup>

$$\Delta\chi_s = E_A^*E_B + C_A^*C_B + W \quad (2)$$

where  $E_A^*$  and  $C_A^*$  measure the influence of electrostatic and covalent bonding respectively on the acceptor property and  $E_B$  and  $C_B$  represent donor tendencies to undergo electrostatic and covalent bonding. Equation 2 is written for an acceptor solute, A, in a series of donor solvents, B. Reported donor parameters<sup>14</sup> are substituted into eq 2 and solved for  $E_A^*$  and  $C_A^*$  to correlate the specific component of the interaction. When both nonspecific and specific interactions contribute to a measured property, eqs 1 and 2 are combined:

$$\Delta\chi = PS' + E_A^*E_B + C_A^*C_B + W \quad (3)$$

The combined equation is solved for  $P$ ,  $E_A^*$ ,  $C_A^*$ , and  $W$  using the solvent dependent  $\Delta\chi$  values and reported<sup>14</sup>  $E_B$ ,  $C_B$ , and  $S'$  values.

Solubility and heats of solution are not treated with eq 3 because they have an added contribution from a free energy or enthalpy of forming a cavity in the solvent to accommodate the solute. Adding such a term to eq 3 could provide all the benefits of the USM model to the problem of solubility prediction. In this report, a correlation of  $S'$  to the solubility parameters was discovered leading to the possibility that  $S'^2$  could provide an estimate of the energy to form the cavity without introducing a new empirical solvent parameter. This new approach to predicting solubilities and enthalpies of solution of gaseous, liquid, and solid solutes not only correlates these data but provides a reliable model to predict these properties for any solvent whose  $S'$  value is known. Furthermore, we will show that the resulting solute parameters provide added insight to the solvation process.

## Calculations

A given data set, consists of solubilities or enthalpies of solution, in kcal mol<sup>-1</sup>, of a given solute in a series of solvents. These quantities are correlated to eq 6 with the NCSS (Kaysville, UT) linear regression program using the reported  $S'$  values and weights summarized in Table 1. The weight in Table 1 is related to the certainty of the values. To arrive at the reported fit, any system (up to a maximum of 3) is removed from the data set if it deviates by 2.5 times the average deviation and the fit rerun. The fit is repeated until all such systems are removed with systems previously removed added back if they subsequently deviate by less than 2.5 times the average deviation.

In some instances a large average deviation results with no system deviating by 2.5 times the average. The system with the largest deviation is removed and the fit rerun. This is repeated. Systems are removed from the final fit only if they deviate by 2.5 times the average deviation. Any system omitted from a data fit is indicated in the tables of results.

In fitting donor-acceptor adduct formation enthalpies measured in nonbasic, poorly solvating solvents to eq 6,  $S'$  is set at zero and  $W$  added. The fit is repeated using the new  $W$  value and reiterated until the  $W$  value does not change.

With  $S'$  values ranging only from 0.6 to 3.1, there is a correlation ( $R^2 = 0.98$ ) between  $S'$  and  $S'^2$ . However, the plot of  $S'^2$  vs  $S'$  is curved, showing positive deviations at the two

ends and negative deviations in the middle. This emphasizes the importance of using solvents that span the  $S'$  range in data fits to eq 6. The correlation of  $S'$  to  $E_B$ ,  $S'$  to  $C_B$ , and  $C_B$  to  $E_B$  for all the solvents in Table 1 give  $R^2$  values of 0.43, 0.01, and 0.09, respectively. In this article we report only systems in which the  $t$  values of  $P$  and/or  $Q$  are significant at least at the 95% confidence level. Even in cases where this is not true, the average deviation of the fit usually is better than experimental error.

The inclusion of nonpolar solvents is critical for polar solutes, but the heat of solution must be measured at low concentration so the solute does not aggregate in solvents of low polarity. In a similar fashion polar solvents are critical to accurate determination of  $P$  and  $Q$  for nonpolar solutes, but the measurements must be made at low solute concentration.

As in solvent polarity fits<sup>18</sup> the solvent [(CH<sub>3</sub>)<sub>2</sub>N]<sub>3</sub>PO deviates in many systems. This solvent and others that contain more than one bond dipole of comparable magnitude in the molecule may orient differently toward different solutes in order to maximize the net solvent-solvent and solvent-solute interaction. Since different solute dependent solvent orientation would each require different values for  $S'$ , data in [(CH<sub>3</sub>)<sub>2</sub>N]<sub>3</sub>PO are not considered in these analyses.

## Results and Discussion

**Relation of  $S'$  to  $\delta$ ,  $\delta_p$ , and  $\delta_d$ .** All nonprotonic donor solvents whose  $S'$  and  $\delta$  values are known were fit to a linear equation in  $\delta$  producing a poor correlation ( $SD = 0.30$ ,  $R^2 = 0.62$  and  $n = 36$ ). Since  $\delta$  correlations apply to nonpolar solvents, the fit of  $\delta$  to  $S'$  was restricted to cyclohexane ( $S' = 1.11$ , wt 1), CCl<sub>4</sub> ( $S' = 1.49$ , wt 0.5), hexane ( $S' = 0.79$ , wt 0.2), heptane ( $S' = 0.9$ , wt 0.05), toluene ( $S' = 1.66$ , wt 0.1), and benzene ( $S' = 1.73$ , wt 0.1), eq 4 results:

$$S' = 0.547(\pm 0.07)\delta - 3.31(\pm 0.6) \quad (4)$$

where  $R^2 = 0.93$  and  $SD = 0.04$ . Though the number of solvents is limited, this is a particularly important result because it provides  $S'$  values for alkanes.<sup>19</sup> It is difficult to obtain  $S'$  for alkanes from spectral probes, which are usually polar, because of aggregation. A given solute molecule is surrounded not by solvent molecules but by a combination of solvent and solute molecules. Until more nonpolar solute spectral probes are developed, tentative  $S'$  values can be more accurately estimated from  $\delta$  than from spectral shifts (see Table 1).

The  $S'$  values of *basic polar solvents* show a good correlation to the Hansen solubility parameters,<sup>6</sup> giving eq 5:

$$S' = 1.228(\pm 0.024) + 0.153(\pm 0.009)\delta_p + 0.062(\pm 0.03)\delta_d \quad (5)$$

$R^2 = 0.91$ ,  $n = 31$ ,  $SD = 0.01$ . This is an interesting result because  $S'$  has its basis in spectral shifts and the Hansen parameters in enthalpies of vaporization and molecular properties.

$S'$  values of polar bases have contributions from both dispersion and polar forces with the *differences* in the  $S'$  values being dominated by dipolar effects.  $S'$  provides a single parameter, nonspecific, solvation scale encompassing solvent dispersion and polar effects that are consistent with the Hildebrand parameters for nonpolar solvents and the Hansen parameters for polar solvents. It can be anticipated that  $S'$  correlations in polar solvents will not apply to a solute property in which polarity contributions are minor and dispersion forces make the main contribution.

TABLE 1: Solvent Parameters Used in These Analyses

solvent	$S'$	$E_B$	$C_B$	wt <sup>d</sup>
pentane <sup>a</sup>	0.57	0.00	0.00	0.2
hexane <sup>a</sup>	0.68	0.00	0.00	0.2
heptane <sup>a</sup>	0.79	0.00	0.00	0.2
nonane <sup>a</sup>	0.90	0.00	0.00	0.2
decane <sup>a</sup>	0.90	0.00	0.00	0.2
C <sub>6</sub> H <sub>12</sub>	1.11	0.00	0.00	1
CCl <sub>4</sub> <sup>c</sup>	1.49	0.00	0.00	1
CS <sub>2</sub>	1.51	0	0	0.2
C <sub>6</sub> H <sub>5</sub> CH <sub>3</sub>	1.66	0.75		0.8
C <sub>6</sub> H <sub>6</sub>	1.73	0.75	0.45	1
C <sub>6</sub> H <sub>5</sub> OCH <sub>3</sub>	2.04	1.20		0.6
C <sub>6</sub> H <sub>5</sub> Cl	2.07	0.70		0.8
C <sub>6</sub> H <sub>5</sub> Br	2.10	0.70		0.4
C <sub>6</sub> H <sub>5</sub> CN	2.63	1.65	0.75	1
C <sub>6</sub> H <sub>4</sub> Cl <sub>2</sub>	2.10	0.30		0.6
C <sub>6</sub> H <sub>5</sub> NO <sub>2</sub>	2.61	1.27	0.57	1
<sup>i</sup> Pr <sub>2</sub> O <sup>b</sup>	1.76	1.75	1.66	0.6
(C <sub>2</sub> H <sub>5</sub> ) <sub>2</sub> O	1.73	1.80	1.63	1
<i>n</i> -bu <sub>2</sub> O	1.58	1.89	1.67	0.8
dioxane	1.93	1.86	1.29	1
C <sub>6</sub> H <sub>5</sub> C(O)CH <sub>3</sub>	2.52	1.72	1.15	0.8
CH <sub>3</sub> COC <sub>2</sub> H <sub>5</sub>	2.50	1.67	1.24	0.8
CH <sub>3</sub> C(O)Ohex <sup>b</sup>	1.94	1.62	0.98	0.4
CH <sub>3</sub> C(O)Obu <sup>b</sup>	1.99	1.62	0.98	0.4
CH <sub>3</sub> C(O)OPr <sup>b</sup>	2.05	1.62	0.98	0.4
CH <sub>3</sub> C(O)OEt	2.15	1.62	0.98	1
HCOOC <sub>2</sub> H <sub>5</sub> <sup>b</sup>	2.24	1.6	0.6	0.4
(CH <sub>3</sub> ) <sub>2</sub> CO	2.58	1.74	1.26	1
C <sub>5</sub> H <sub>5</sub> N	2.44	1.78	3.54	1
4-CH <sub>3</sub> C <sub>5</sub> H <sub>4</sub> N	2.31	1.83	3.73	0.8
HC(O)NEt <sub>2</sub> <sup>b</sup>	2.59	1.31	2.19	0.4
CH <sub>3</sub> C(O)N(CH <sub>3</sub> ) <sub>2</sub>	2.70	2.35	1.31	1
HC(O)N(CH <sub>3</sub> ) <sub>2</sub>	2.80	2.19	1.31	1
[(CH <sub>3</sub> ) <sub>2</sub> N] <sub>2</sub> CO	2.55	2.06	1.87	0.6
(CH <sub>3</sub> ) <sub>2</sub> SO	3.00	2.40	1.47	1
(CH <sub>2</sub> ) <sub>4</sub> O	2.08	1.64	2.18	1
OP[N(CH <sub>3</sub> ) <sub>2</sub> ] <sub>3</sub>	2.52	2.87	1.52	0.6
C <sub>2</sub> H <sub>5</sub> CN	2.80	1.40		0.6
<i>n</i> -C <sub>3</sub> H <sub>7</sub> CN	2.70	1.81	0.54	0.6
CH <sub>3</sub> CN	3.00	1.64	0.71	1
CH <sub>3</sub> NO <sub>2</sub>	3.07	1.09	0.70	0.8
(C <sub>2</sub> H <sub>5</sub> ) <sub>3</sub> N	1.43	1.32	5.73	1
MIm <sup>e</sup>	2.6	1.16	4.92	0.8
PC <sup>e</sup>	3.10	1.51		0.6
MeSul <sup>e</sup>	2.55	1.61	1.09	0.4
NMP	2.62	2.97		0.6

<sup>a</sup> Calculated with eq 4. If the measurement is not carried out in very dilute solution  $<5 \times 10^{-5}$  a weight of 0.01 or less is used.

<sup>b</sup> Calculated with eq 5. <sup>c</sup> Specific interactions are a problem. <sup>d</sup> Weights are determined by subtracting 0.4 from one if both  $S'$  and  $E_B$  are tentative and 0.2 if either one is. <sup>e</sup> MIm refers to *N*-methylimidazole, PC to propylene carbonate, MeSul to 3-methylsulfolane, and NMP to *N*-methylpyrrolidinone.

**Cavity Term.** The square of the solubility parameter,  $\delta^2$ , provides the energy to overcome solvent cohesion forces to create a hole for the solute in nonpolar solvents. The relation of  $S'$  to  $\delta$  suggests that  $S'^2$  would provide a comparable measure of cavity formation in both polar and nonpolar solvents that are not associated via specific hydrogen-bonding interactions. Adding a cavity term  $QS'^2$  to eq 3, where  $Q$  is a solute-dependent parameter, leads to the unified solvation cavity model (USCM) given by eq 6:

$$\Delta\chi = QS'^2 + PS' + E_A^*E_B + C_A^*C_B + W \quad (6)$$

$Q$ ,  $P$ ,  $E_A^*$ ,  $C_A^*$ , and  $W$  are the solute parameters defined above.  $W$  gives all the contributions to  $\Delta\chi$  that exist in a solvent for which  $S' = E_B = C_B = 0$  including the energy to separate a given liquid or solid solute or condense a gaseous solute to form

the solution molecular species. When a protic solute is studied in a series of nonpolar and donor solvents, the solute–solute hydrogen-bond dissociation energy is constant and becomes incorporated into  $W$ . By analogy to scaled particle theory, the addition of an  $S'^2$  term can be considered to convert eq 3 into a virial equation with cavity formation represented by  $QS'^2$  and solute–solvent interaction represented by  $PS'$ .

**Enthalpy of Solution of Donor Solutes in Nonpolar and Polar Donor Solvents.** The first systems that will be discussed involve the heat of solution of donor and hydrocarbon solute molecules in nonacidic solvents.<sup>20–22</sup> The donor solute, *N*-methylimidazole (MIm) has been studied<sup>22a</sup> in 24 nonacidic solvents whose  $E_B$ ,  $C_B$ , and  $S'$  values are known. The data and fit of this liquid solute, given in Table 2, will be discussed in detail for these are ideal measurements on dilute solutions that span the  $S'$  range.

The data fit of *N*-methylimidazole to eq 6 was attempted with the  $E_AE_B$  and  $C_AC_B$  terms set at zero because specific donor–acceptor interactions are not anticipated for a donor solute in donor solvents. A very good correlation resulted ( $R^2 = 0.955$ ). However, the fit was improved significantly ( $R^2 = 0.98$ ) by adding the  $E_A^*E_B$  term of eq 6. Although  $E_B$  contributes only 2% to  $R^2$ , the  $t$  value indicates it is statistically significant at the 99.5% confidence level. The  $C_AC_B$  term is not significant.

The need for an  $E_A^*E_B$  term on a donor solute that is not expected to undergo specific donor–acceptor interactions with donor solvents is a surprising and significant result. With nonspecific solvation, the solute is oriented in the solvent cavity to maximize the nonspecific dipolar and dispersion interactions of the cavity and the solute. The enthalpy of stabilization of *N*-methylimidazole in dimethyl sulfoxide from this  $PS'$  effect is about  $-24 \text{ kcal mol}^{-1}$ . If the donor solute is so oriented that it can behave as a weak acceptor toward a solvent lone pair, a weak, electrostatic, specific interaction can occur whose enthalpy is given by  $E_A^*E_B$ . This is a small enthalpy term that produces only  $-0.5 \text{ kcal mol}^{-1}$  for *N*-methylimidazole dissolved in (CH<sub>3</sub>)<sub>2</sub>SO. In the absence of the  $-24 \text{ kcal mol}^{-1}$  orienting effect of nonspecific solvation, the specific interaction would lead to too low an equilibrium constant for the adduct to form in observable amounts. The average orientation of the solute with respect to the solvent is one that maximizes the cavity–solute and solvent lone pair–solute interactions.

The measured enthalpies for MIm in benzene and toluene are  $\sim 0.8 \text{ kcal mol}^{-1}$  more exothermic than the calculated values. Considering all possible reasons for these deviations in the context of patterns observed in all the fits reported, we conclude that MIm aggregates in benzene and toluene. The  $S'^2$  values indicate that it is twice as hard to form cavities in benzene and toluene as it is in cyclohexane. Deviations in the same direction are observed for *N*-methylpyrrole, but their magnitude is not sufficient to remove them from the fit.

The enthalpy of solution of MIm in CCl<sub>4</sub> deviates from the model. It was shown earlier<sup>23</sup> that enthalpies of adduct formation measured in CCl<sub>4</sub> with strong nitrogen and sulfur donors have contributions from specific interactions of  $\sim 1 \text{ kcal mol}^{-1}$  with CCl<sub>4</sub> behaving as an acceptor solvent. The CCl<sub>4</sub> result is not an exception to the USCM but is omitted from the fit because the  $E_A$  and  $C_A$  parameters for CCl<sub>4</sub> are not known. As expected, solutes whose  $E_B$  and  $C_B$  values are small behave as predicted in CCl<sub>4</sub> while solutes with appreciable values consistently deviate.

Nitromethane was omitted from the MIm data fit and is proposed to hydrogen bond to this solute. Nitromethane is well behaved in the fits of the weaker donor solutes<sup>24</sup> in Tables 2 and 3, consistent with the expected very weak acceptor

TABLE 2: Enthalpies of Solution at Infinite Dilution

solvent <sup>a</sup>	MIm		N-MePyr		N-MePyrz		C <sub>6</sub> H <sub>6</sub>		n-C <sub>4</sub> H <sub>9</sub> OH	
	exp	cal <sup>a</sup>	exp	cal <sup>b</sup>	exp	cal <sup>c</sup>	exp	cal <sup>d</sup>	exp	cal <sup>e</sup>
C <sub>6</sub> H <sub>12</sub>	3.78	3.76	1.89	1.83	4.65	4.62	0.91	0.89		
C <sub>6</sub> H <sub>5</sub> CH <sub>3</sub>	0.90	(1.57)	0.22	0.48			0.06	0.16		
C <sub>6</sub> H <sub>6</sub>	0.85	(1.39)	0.16	0.38			0	0.12	3.93	4.00
C <sub>6</sub> H <sub>5</sub> OCH <sub>3</sub>	0.43	0.62	0.03	-0.02	0.63	0.60	0.05	0.05		
C <sub>6</sub> H <sub>5</sub> Cl	0.50	0.67	-0.02	0.00	0.83	0.84	-0.01	0.02		
(CH <sub>2</sub> ) <sub>4</sub> O	0.43	0.47	-0.20	-0.08	0.10	0.27	-0.34	-0.11	1.06	1.16
o-C <sub>6</sub> H <sub>4</sub> Cl <sub>2</sub>	0.65	0.71	0.05	0.01	0.99	1.02	0.12	0.08		
CH <sub>3</sub> C(O)OEt	0.44	0.37	0.06	-0.12	0.18	0.17	0.09	-0.11	2.13	2.07
CH <sub>3</sub> C(O)C <sub>2</sub> H <sub>5</sub>	0.31	0.08	0.00	-0.16					1.50	1.92
C <sub>5</sub> H <sub>5</sub> N	0.17	0.08	-0.06	-0.19			-0.04	-0.03	0.30	0.08
C <sub>6</sub> H <sub>5</sub> CN	0.15	0.08	-0.20	-0.10	0.30	-0.09	0.05	0.13		
C <sub>6</sub> H <sub>5</sub> NO <sub>2</sub>	0.29	0.15	0.02	-0.08			0.28	0.17		
(CH <sub>3</sub> ) <sub>2</sub> CO	0.23	0.06	0.00	-0.13	-0.21	-0.16				
OPN(CH <sub>3</sub> ) <sub>2</sub> ] <sub>3</sub>	-0.80	(-0.16)	-1.23	(-0.25)	-1.13	(-0.82)	-0.60	(-0.14)		
[(CH <sub>3</sub> ) <sub>2</sub> N] <sub>2</sub> CO	-0.23	0.00	-0.43	(-0.17)	-0.57	-0.35	-0.34	(0.00)		
C <sub>2</sub> H <sub>5</sub> CN	0.32	0.20	-0.03	0.06	0.29	0.17				
(CH <sub>3</sub> ) <sub>2</sub> SO	-0.04	0.19	0.13	0.24	-0.14	-0.13	0.63*	0.46	0.93	0.58
CH <sub>3</sub> CN	0.43	0.35	0.38	0.30			0.62*	0.57	2.52	2.53
CH <sub>3</sub> NO <sub>2</sub>	-0.33	(0.55)	0.52	0.46	0.59	0.77	1.02*	0.76	3.66	3.55
HC(O)N(CH <sub>3</sub> ) <sub>2</sub>	-0.17	0.04	-0.19	0.00	-0.29	-0.28	0.06*	0.23	0.74	1.01
O(C <sub>2</sub> H <sub>4</sub> ) <sub>2</sub> O	0.50	0.69	-0.05	0.03	0.46	0.44	0.05	-0.13	1.72	1.33
(C <sub>2</sub> H <sub>5</sub> ) <sub>3</sub> N	2.18	2.20	0.87	0.89			0.27	0.27	-1.14	-0.98

<sup>a</sup> Calculated for *N*-methylimidazole (MIm) with  $1.55(\pm 0.1)S^2 - 7.98(\pm 0.4)S' - 0.20(\pm 0.07)E_B + 10.71(\pm 0.5)$ ;  $R^2 = 0.97$  and an  $F$  ratio of 314;  $F(0.995, 3, 20) = 2.84$ . Also included: butyl acetate, exp/cal is 0.41/0.62; *n*-Bu<sub>2</sub>O 1.84/1.58. CCl<sub>4</sub> omitted. Reference 22a. The  $t$  values for  $Q$ ,  $P$ ,  $E_A^*$ , and  $W$  are 15, -18, -4.04, and 23;  $t(0.999, 20) = 3.849$ . <sup>b</sup> Calculated for *N*-methylpyrrole with  $1.19(\pm 0.1)S^2 - 5.64(\pm 0.5)S' - 0.08(\pm 0.06)E_B + 6.62(\pm 0.5)$ ;  $R^2 = 0.91$ ,  $F$  ratio is 83;  $F(0.999, 3, 20) = 3.79$ . Also included: *n*-butyl acetate 0.09/0.03; CS<sub>2</sub> 0.95/0.82; *n*-Bu<sub>2</sub>O 0.83/0.53; *N*-methylimidazole 0.02/-0.05. CCl<sub>4</sub> omitted. Reference 22. When differences exist, ref 22a data are used. The  $t$  values for  $Q$ ,  $P$ ,  $E_A^*$ , and  $W$  are 12, -12, -1, 14 [ $t(0.999, 22) = 3.79$ ]. The  $t$  value of  $E_A$  is not significant at 95% confidence. <sup>c</sup> Calculated for *N*-methylpyrazole with  $1.87(\pm 0.1)S^2 - 9.45(\pm 0.6)S' - 0.56(\pm 0.1)E_B + 12.80(\pm 0.6)$ ;  $R^2 = 0.99$ ,  $F$  ratio of 271;  $F(0.999, 3, 10) = 12.6$ . The  $t$  values for  $Q$ ,  $P$ ,  $E_A^*$ , and  $W$  are 21, 14, -16, and -6,  $t(0.999, 11) = 4.44$ . Reference 22. <sup>d</sup> Calculated for benzene with  $0.81 \pm 0.11S^2 - 3.37(\pm 0.5)S' - 0.14(\pm 0.08)E_B + 3.63(\pm 0.5)$ ;  $R^2 = 0.78$ ,  $F$  ratio = 24,  $F(0.999, 3, 15) = 9.34$   $t$  values for  $Q$ ,  $P$ ,  $E_A^*$ , and  $W$  are 6.8, 7.1, -6.5, and -2.0,  $t(0.95, 19) = 1.73$ . Also included: MIm 0.28/0.17; 3-CH<sub>3</sub> sulfolane -0.02/0.03; CH<sub>3</sub>CON(CH<sub>3</sub>)<sub>2</sub> -0.05\*/0.1; CS<sub>2</sub> 0.62/0.39. Asterisks indicate data from ref 20. Other data are from ref 22b. <sup>e</sup> Calculated for *n*-butanol with  $-0.01S^2 + 0.42S' - 1.83E_B - 0.70C_B + 4.91$ ;  $R^2 = 0.97$ ;  $F$  ratio = 81;  $F(0.999, 3, 10) = 12.6$ . The  $t$  values for  $Q$ ,  $P$ ,  $E_A^*$ ,  $C_A^*$ , and  $W$  are -0.04, 0.26, 8.9, 11.6, and 2.8;  $t(0.99, 10) = 2.76$ .  $Q$  and  $P$  are not significant. In addition to the data shown, the following data were included exp/cal: DMA 0.28/0.68; (C<sub>2</sub>H<sub>5</sub>)<sub>2</sub>O 1.15/1.13; 4-CH<sub>3</sub>-C<sub>5</sub>H<sub>4</sub>N -0.09/-0.19. References 21 and 29.

TABLE 3: Parameters for Predicting Enthalpies of Solution at Infinite Dilution for Polar and Nonpolar Probes in Basic Solvents

probe	$Q$	$P$	$E^*$	$W$	$n(n')$	$R^2$	$F^e$	ref
(CH <sub>3</sub> ) <sub>2</sub> SO	$1.48 \pm 0.3$	$-8.02 \pm 1.5$	$-0.05 \pm 0.3^b$	$10.91 \pm 1.5$	9(1)	0.94	36	24
CH <sub>3</sub> C(O)OC <sub>2</sub> H <sub>5</sub>	$1.14 \pm 0.2$	$-5.77 \pm 0.8$	$0.42 \pm 0.2$	$6.52 \pm 0.9$	10(0)	0.89	19.6	24
<i>n</i> -C <sub>4</sub> H <sub>9</sub> Cl <sup>g</sup>	$0.79 \pm 0.2$	$-2.70 \pm 0.9$	$-0.17 \pm 0.1^b$	$2.42 \pm 0.9$	14(1)	0.83	18	29
CH <sub>3</sub> CN <sup>g</sup>	$0.76 \pm 0.2$	$-4.01 \pm 1$	$-0.33 \pm 0.15$	$5.94 \pm 1.3$	9(0)	0.96	28	24
C <sub>6</sub> H <sub>5</sub> CH <sub>3</sub>	$0.91 \pm 0.1$	$-3.64 \pm 0.6$	$-0.14 \pm 0.1^b$	$3.75 \pm 0.6$	21(0)	0.80	23	22a, 21
C <sub>6</sub> H <sub>5</sub> OCH <sub>3</sub>	$1.35 \pm 0.2$	$-6.09 \pm 0.7$	$-0.15 \pm 0.1^b$	$6.53 \pm 0.7$	17(1)	0.87	31	29
4-FC <sub>6</sub> H <sub>4</sub> OCH <sub>3</sub>	$1.36 \pm 0.2$	$-5.92 \pm 0.8$	$-0.54 \pm 0.2$	$6.80 \pm 0.8$	17(3)	0.94	76	29
<i>n</i> -C <sub>6</sub> H <sub>14</sub> (g)	$0.88 \pm 0.3$	$-2.36 \pm 1.2$	$0.10 \pm 0.2^b$	$-5.84 \pm 1$	12(0)	0.91	29	13
(CH <sub>3</sub> ) <sub>4</sub> Sn(g)	$0.58 \pm 0.2$	$-1.73 \pm 0.9$	$0.34 \pm 0.2$	$-6.17 \pm 0.9$	11(0)	0.91	28	13
C <sub>6</sub> H <sub>12</sub> (g)	$0.36 \pm 0.2$	$-0.86 \pm 0.7^b$	$0.51 \pm 0.2$	$-7.27 \pm 0.7$	10(0)	0.96	64	13
nitrocellulose	$0.18 \pm 0.7^b$	$-0.81 \pm 3$	$-1.51 \pm 0.2$	$3.82 \pm 3^b$	6(0)	0.99	111	24
<i>t</i> -BuOH <sup>c</sup>	$1.58 \pm 0.1$	$-7.69 \pm 0.2$	$c$	$13.22 \pm 0.8$	10(1)	0.99	772	29

<sup>a</sup> All probes are in their physical state, *i.e.* liquid or solid, at 25 °C unless indicated, (g) refers to gas.  $n$  refers to the number of data points used in the fit and ( $n'$ ) to the number of solvents omitted. [(CH<sub>3</sub>)<sub>2</sub>N]<sub>3</sub>PO is omitted from all data fits and not counted for the  $n'$  value. See text. <sup>b</sup> The need for this term is not statistically significant at 95% confidence for the solvent set used. <sup>c</sup> Heat of solution of *tert*-butyl alcohol. The following solvation minimized enthalpies were added (see calculation section) C<sub>5</sub>H<sub>5</sub>N -4.30, (CH<sub>3</sub>)<sub>2</sub>SO -3.60, 4-CH<sub>3</sub>-C<sub>5</sub>H<sub>4</sub>N -4.6. The heat of solution in pyridine was omitted. The fit produces values of  $E_A = -1.09 \pm 0.3$  and  $C_A = -0.68 \pm 0.07$ . <sup>d</sup> The enthalpy of solution of nitrocellulose. No data were measured in hydrocarbon solvents. <sup>e</sup> All  $F$  values are significant at the 99.5% confidence level.

properties of this solvent. The above discussion illustrates the detailed understanding about the solution process that can result from analyses with eq 6.

The enthalpies of solution of liquid *N*-methylpyrrole (*N*-MePyr of Table 2), studied<sup>22a</sup> in 23 solvents whose  $E_B$  and  $S'$  values are known, give the fit to eq 6 shown in Table 2. Small deviations in the same direction as *N*-methylimidazole are observed in benzene and toluene. The fit of this weakly basic solute in CH<sub>3</sub>NO<sub>2</sub> supports the interpretation of the donor-

acceptor interaction between CH<sub>3</sub>NO<sub>2</sub> and the stronger donor *N*-methylimidazole.

The parameters from the data fits of other liquid solutes, are reported in Table 3. The correlation to eq 6 is excellent for both polar and nonpolar solutes dissolved in both polar and nonpolar solvents giving signs for the solute parameters that are consistent with the enthalpic contribution expected for the terms from the model. Compared to  $PS'$ , the  $E_A^*E_B$  term makes a small contribution ( $<1$  kcal mol<sup>-1</sup>) to solute-solvent interac-

tion which, in some cases, is statistically nonsignificant. Positive  $E_A^*$  values, which would suggest an antibonding solvent–solute lone-pair repulsion, are obtained for some solutes, but the parameters are zero within experimental error.

The enthalpies of solution<sup>13</sup> of the gaseous nonpolar solutes  $n$ -C<sub>6</sub>H<sub>14</sub>,  $c$ -C<sub>6</sub>H<sub>12</sub> and (CH<sub>3</sub>)<sub>4</sub>Sn lead to negative values for  $W$  corresponding to the exotherm expected<sup>2</sup> for adding a gas to a solvent. Adding the heat of vaporization of the solute to the enthalpies of solution of the gas solute, led to experimental enthalpies of solution of the liquid solute. As expected from the model, the only changes in the parameters for fits of the same gaseous and liquid solute should be and are in the  $W$  value. Note for cyclohexane, the negative  $W$  for the gaseous solute and the positive  $W$  for the liquid.

The data fits for the solutes, toluene, anisole, and 4-fluorani-sole (Table 3), are similar to those shown for benzene as a solute in Table 2. The benzene fit produces an  $R^2$  of 0.86. A similar result is obtained for toluene (Table 3). The small range of measured enthalpies leads to a low  $R^2$  with good residuals (0.1–0.3 kcal mol<sup>-1</sup>).

The signs of the  $P$ ,  $Q$ , and  $W$  parameters for all the data fits indicate that the correlations are not meaningless, but they support the USCM interpretation. A consistent pattern of endoergic contributions for  $Q$  (cavity formation) and exoergic contributions from  $P$ ,  $E^*$ , and  $C_A^*$  (nonspecific and specific solvation) is found for all solute data sets. The endoergic  $W$  found for liquid and solid solutes as well as the exoergic  $W$  found for gases also is expected.

**Enthalpies of Solution of Acceptor Solutes in Donor Solvents.** When acceptor solutes are added to donor solvents, specific donor–acceptor interactions occur forming an adduct that is nonspecifically solvated. The data fit for the enthalpy of solution of  $n$ -butanol in different solvents to eq 6 is excellent as shown in Table 2. The resulting  $E_A$  and  $C_A$  values are reasonable when compared to  $E_A$  and  $C_A$  for aliphatic alcohols. Since these terms refer to the exothermic donor–acceptor contribution to the heat of solution, the negative signs are expected. The  $E$  and  $C$  parameters are positive in the ECW model (eq 2) because  $-\Delta H$  is fit. Careful attention has to be paid to the property being fit to interpret the signs of the parameters. Because data in nonbasic, low  $S'$  solvents are absent,  $t$  values of 0.26 and  $-0.04$  result for  $P$  and  $Q$ , respectively, suggesting these terms are not significant at 95% confidence. The specific interaction dominates the solution enthalpy differences for  $n$ -butanol in different solvents.

Only limited data is available for *tert*-butyl alcohol, and there is none in low- $S'$  solvents. The data are fit (Table 3) by combining the heats of solution with adduct formation enthalpies measured in poorly solvating solvents (see Calculation section). The  $E_A$  and  $C_A$  values of 1.15 and 0.66 are similar to those for  $n$ -C<sub>4</sub>H<sub>9</sub>OH and are within error of the reported<sup>14</sup> tentative values of 1.07 and 0.69 for this acceptor.

**USM Parameters and View of Cavity Formation.** The impressive correlation of enthalpy data with the USM parameters and USCM (eq 6) lends confidence in the ability of the model to predict enthalpies of new systems. Of even greater significance are the insights about the solution process contained in the parameters. The eventual worth of the model will be measured by the insights provided by future measurements and the experiments suggested to test the model. Even at this early stage, the set of parameters from the enthalpy data fits presented in Tables 2 and 3 lead to a proposal for a revised view of cavity formation. The liquid state can be viewed as a dynamic distribution of various size clusters in which solvent molecules are held together in orientations similar to those in the

corresponding solid. When a solute is added, it replaces a solvent molecule in a cluster and orients so as to maximize the dispersion and polar interactions given by  $PS'$ . Because of the dynamic nature of the system, average energies result. The incorporation of a solute molecule changes the geometry of the solvent cluster particularly in the region surrounding the solute. The distribution of the sizes of the solvent clusters can also change. The geometry change in the solvent component is endothermic and makes a positive contribution to the free energy. Change in cluster size is expected to be largely entropic, making the free energy of solution more negative if the average size decreases and making the free energy more positive if the average size increases. These solvent terms constitute a new USCM interpretation of cavity formation. The resistance to inducing these changes in a solvent is given by  $S'^2$  and the extent to which a solute induces these changes is given by  $Q$ .

The  $Q$  values of  $\sim 0.75 \pm 0.1$  for nonpolar solutes ( $n$ -C<sub>4</sub>H<sub>9</sub>-Cl, C<sub>6</sub>H<sub>6</sub>, C<sub>6</sub>H<sub>5</sub>CH<sub>3</sub>,  $n$ -C<sub>6</sub>H<sub>14</sub>, (CH<sub>3</sub>)<sub>4</sub>Sn) suggest, in contrast to perceptions from other models,<sup>2,6,13</sup> that the variation in the molecular size and shape of nonpolar solutes has only a minor influence on the energy of cavity formation. All these different size nonpolar solutes have similar  $Q$  values inducing similar energy cavity ( $QS'^2$ ) changes. The larger  $Q$  found for non-hydrogen-bonding, polar solutes than for nonpolar solutes corresponds to a greater distortion in the neighboring solvent–solvent interactions and cluster sizes to maximize the  $PS'$  interaction. This is a picture of cavity formation very different from that of other models<sup>2,6,10</sup> where cavity formation is viewed to be more endothermic for larger solutes.

The most compelling argument for this dynamic cluster model of cavity formation comes from the  $Q$  values of the solutes  $t$ -C<sub>4</sub>H<sub>9</sub>OH and  $n$ -C<sub>4</sub>H<sub>9</sub>OH. When an acceptor molecule is added to a series of basic solvents, each base molecule has a different size and different size adducts would form. The value of  $Q$  would vary for each basic solvent and eq 6 would not work. The constant value of  $Q$  in USCM suggests that  $B$  remains an integral part of the solvent cluster and the alcohols hydrogen bond to a base in a cluster of base molecules.

**Other Properties Requiring a Cavity Term.** Up to this point all systems discussed involve enthalpies. One very important application of regular solution theory is the prediction of solubility,<sup>2,6</sup> a free energy. The  $E_B$  and  $C_B$  parameters have been used<sup>14</sup> to correlate enthalpies, free energies, and entropies. As with Hammett type linear free energy correlations, conditions necessary for free energy and entropy correlations to ECW have been reported.<sup>25</sup> The application of eq 6 to solubility prediction is demonstrated with the free energy fits (kcal mol<sup>-1</sup>) in Tables 4 and 5. The solubilities of a series of solid benzoic acid derivatives<sup>5c,d</sup> in different solvents (Table 4), expressed as the natural logs of solubility mole fraction are fit to eq 6. The signs on the solute fit parameters (Table 4) are in agreement with the USCM prediction that positive signs on  $E_A^*$ ,  $C_A^*$ , and  $P$  decrease log solubility, increasing solubility, whereas the negative  $Q$  for the cavity term decreases solubility. The free energy fit produces parameters of opposite sign.

The solute-dependent intercept,  $W$ , has as a main component the free energy needed to overcome the intermolecular forces of the solid or liquid solute to disperse it in the solvent. The relative increasing negative values of the intercepts ( $-5.87$ ,  $-15.39$ , and  $-38.36$ ) of the solids benzoic acid (BA), methyl *p*-hydroxybenzoate (MB), and *p*-hydroxybenzoic acid (PB) are highly correlated to the corresponding increase of negative  $\ln X_2^f$  values<sup>26</sup> of solubility parameter theory.

The correlations of free energies of solution<sup>27</sup> of the intimate ion pair Et<sub>4</sub>N<sup>+</sup>I<sup>-</sup>, the solvent-separated ion pair Et<sub>4</sub>N<sup>+</sup>, I<sup>-</sup> and

TABLE 4: Free Energies of Solution and  $\ln$  Solubilities of Solutes

solvents	benzoic acid		Me- <i>p</i> -hydroxybenzoate		<i>n</i> -octane		nitromethane		Et <sub>4</sub> N <sup>+</sup> I <sup>-</sup>	
	$\ln X_2 \text{ exp}^a$	$\ln X_2 \text{ cal}^b$	$\ln X_2 \text{ exp}^a$	$\ln X_2 \text{ cal}^c$	$\Delta G \text{ exp}^d$	$\Delta G \text{ cal}^e$	$\Delta G \text{ exp}^d$	$\Delta G \text{ cal}^f$	$\Delta G \text{ exp}^g$	$\Delta G \text{ cal}^h$
pentane	-5.13	-4.98	-9.67	-10.31						
hexane	-4.66	-4.87	-9.47	-9.64	-2.42	-2.43	0.19	(1.0)		
heptane	-4.45	-4.74	-9.40	-8.86						
nonane	-4.26	-4.62	-9.15	-8.12						
decane	-4.17	-4.62	-9.02	-8.12	-2.53	-2.37				
C <sub>6</sub> H <sub>12</sub>	-4.59	-4.41			-2.42	-2.46	0.35	0.14		
C <sub>6</sub> H <sub>5</sub> CH <sub>3</sub>	-2.66	-3.01			-2.16	-2.23				
C <sub>6</sub> H <sub>6</sub>	-2.67	-2.98			-2.02	-2.18				
C <sub>6</sub> H <sub>5</sub> Cl	-2.45	-2.58			-1.90	-1.80			6.39	6.87
C <sub>6</sub> H <sub>5</sub> NO <sub>2</sub>	-2.41	-2.19			-0.93	-0.79			3.40	3.51
(CH <sub>2</sub> ) <sub>4</sub> O					-2.00	-1.74	-1.71	-1.53		
(C <sub>2</sub> H <sub>5</sub> ) <sub>2</sub> O	-1.69	-1.41	-2.48	-2.94			-0.95	-1.13	10.15	10.17
CH <sub>3</sub> NO <sub>2</sub>					0.47	0.41			2.51	2.45
C <sub>6</sub> H <sub>5</sub> C(O)CH <sub>3</sub>	-1.67	-1.49							3.98	3.91
CH <sub>3</sub> C(O)Obu	-1.77	-1.41	-2.02	-2.19						
CH <sub>3</sub> C(O)OEt	-1.80	-1.69	-2.07	-1.77	-1.44	-1.64	-1.80	-1.60	6.63	6.23
(CH <sub>3</sub> ) <sub>2</sub> CO	-1.68	-1.44			-0.93	-0.83	-1.94	-1.92	3.57	3.64
C <sub>5</sub> H <sub>5</sub> N	-0.63	-0.70	-1.13	-0.97						
CH <sub>3</sub> C(O)N(CH <sub>3</sub> ) <sub>2</sub>	-0.65	-0.78	-0.61	-0.64						
HC(O)N(CH <sub>3</sub> ) <sub>2</sub>	-0.71	-0.95	-0.77	-0.66	-0.36	-0.29	-2.25	-2.19	2.21	2.87
(CH <sub>3</sub> ) <sub>2</sub> SO	-0.67	-0.72	-0.54	-0.58	0.62	0.27	-2.19	-2.31		
CH <sub>3</sub> CN					0.06	0.23			2.65	2.45

<sup>a</sup> Reference 5c,d. <sup>b</sup> Calculated for benzoic acid with  $\ln X_2 = -0.352(\pm 0.2)S'^2 + 1.70(\pm 0.9)S' + 1.15(\pm 0.4)E_B + 0.30(\pm 0.1)C_B - 5.87(\pm 0.8)$ ;  $r^2 = 0.96$ , SD = 0.20,  $n = 21$ . Also included, dioxane  $-1.25/-1.36$ . The corresponding equation for the free energy of solution of benzoic acid is  $\Delta G = 0.21(\pm 0.1)S'^2 - 1.01(\pm 0.5)S' - 0.68(\pm 0.1)E_B - 0.18(\pm 0.05)C_B + 3.47(\pm 0.4)$ . The  $t$  values for  $Q$ ,  $P$ ,  $E_A^*$ ,  $C_A^*$ , and  $W$  are 2.3, 2.7, 6.7, 3.5, and -11;  $t(0.98,12) = 2.22$ .  $R^2 = 0.96$ , SD = 0.1,  $F = 129$ ;  $F(0.999;4,15) = 8.25$ . <sup>c</sup> Calculated for methyl *p*-hydroxybenzoate with  $\ln X_2 = -1.68(\pm 0.3)S'^2 + 9.58(\pm 1.4)S' + 0.426(\pm 0.2)E_B + 0.07(\pm 0.1)C_B - 15.39(\pm 1.5)$ ;  $r^2 = 0.95$ , SD = 0.2. Also included in the fit (*n*-C<sub>4</sub>H<sub>9</sub>)<sub>2</sub>O  $3.62/3.55$ ; CH<sub>3</sub>COOC<sub>3</sub>H<sub>7</sub>  $-1.99/-2.02$ ; HC(O)N(C<sub>2</sub>H<sub>5</sub>)<sub>2</sub>  $-0.71/-1.09$ , CH<sub>3</sub>COOhex  $-2.15/-2.34$ , <sup>i</sup>Pr<sub>2</sub>O  $-3.46/-2.89$ . The corresponding equation for the free energy is  $\Delta G = 0.99(\pm 0.2)S'^2 - 5.67(\pm 0.8)S' - 0.25(\pm 0.1)E_B - 0.04(\pm 0.05)C_B + 9.11(\pm 0.8)$ . The  $t$  values are -6, 8, 2, 1, 17;  $t(0.95,12) = 1.78$ .  $R^2 = 0.95$ , SD = 0.1,  $F = 149$ ;  $F(0.999;4,15) = 8.25$ . <sup>d</sup> Reference 13. <sup>e</sup> Free energy of solution of *n*-octane in solvents. Calculated with  $0.80(\pm 0.13)S'^2 - 1.91(\pm 0.6)S' + 0.05E_B(\pm 0.1) - 1.29(\pm 0.6)$ .  $t$  values are 5.9, 3.2, 0.5, and 2.2;  $t(0.95,14) = 1.76$ .  $R^2 = 0.96$ ,  $F = 138$ ,  $n = 19$ ;  $F(0.999;3,12) = 10.8$ . Also included (C<sub>2</sub>H<sub>5</sub>)<sub>3</sub>N  $-2.56/-2.33$ ; PC/0.32/0.50; NMP  $-0.35/-0.68$ ; CCl<sub>4</sub>  $1.97/2.37$ ; (*n*-C<sub>4</sub>H<sub>9</sub>)<sub>2</sub>O  $-2.39/-2.23$ ; C<sub>6</sub>H<sub>5</sub>Br  $-1.95/-1.76$ . <sup>f</sup> Free energy of solution of CH<sub>3</sub>NO<sub>2</sub>. Calculated with  $0.70(\pm 0.13)S'^2 - 4.15(\pm 0.60)S' - 0.06(\pm 0.13)E_B + 4.07$ .  $t = 5.3$ , -6.8, 0.5, 7.4.  $R^2 = 0.99$ , SD = 0.11,  $F = 155$ . Also included (C<sub>2</sub>H<sub>5</sub>)<sub>3</sub>N  $-0.57/-0.62$ ; CH<sub>3</sub>C(O)C<sub>2</sub>H<sub>5</sub>  $-1.92/-1.92$ . *n*-hexane omitted. <sup>g</sup> Reference 27. <sup>h</sup> Free energy of solution of Et<sub>4</sub>N<sup>+</sup>I<sup>-</sup> intimate ion pair. Calculated with  $3.93(\pm 0.6)S'^2 - 24.60(\pm 3.0)S' + 40.98(\pm 3.62)$ . The  $t$  values are 6.4, -8.2, 11.3.  $R^2 = 0.98$ , SD = 0.35,  $F = 223$ .

TABLE 5: Free Energies and Enthalpies of Solution of Solutes

	$Q$	$P$	$E_A$	$C_A$	$W$	$n$	$R^2$	$F$	ref
$\ln X_2 p\text{-HOC}_6\text{H}_4\text{COOH}$	-5.09(±1.7)	22.55(±7)	6.12(±0.8)	0.18(±0.3)	-38.36(±1.5)	13	0.95	149	5d
$\Delta G p\text{-HOC}_6\text{H}_4\text{COOH}$	3.01(±0.9)	-13.35(±4)	-3.63(±0.5)	-0.10(±0.2)	22.71(±5)	13	0.95	149	5d
$\Delta H \text{ poly(D,L-lactide)}^a$	2.33(±0.2)	-13.77(±1)	-3.63(±0.2)	-0.78(±0.1)	25.42(±1)	6	0.99 <sup>a</sup>	235	28
$\log LE^b$	0.29(±0.03)	-2.09(±0.1)	-0.80(±0.02)	-0.28(±0.01)	6.57(±0.2)	6	0.999	1013	28
$\Delta G \text{ C}_2\text{H}_6(\text{g})^d$	0.29(±0.1)	-0.59(±0.3) <sup>e</sup>	-0.02(±0.05) <sup>e</sup>		2.44(±0.3)	13	0.93	50	13
$\Delta G \text{ C}_3\text{H}_8(\text{g})^d$	0.45(±0.1)	-1.18(±0.4)	0.12(±0.1) <sup>e</sup>		2.10(±0.4)	12	0.95	56	13
$\Delta G n\text{-C}_4\text{H}_{10}(\text{g})^d$	0.42(±0.1)	-1.01(±0.4)	0.16(±0.07)		1.12(±0.4)	9	0.98	134	13
$\Delta G n\text{-C}_5\text{H}_{12}(\text{g})^d$	0.70(±0.2)	-1.95(±0.9)	0.08(±0.1) <sup>e</sup>		1.16(±0.8) <sup>e</sup>	13	0.91	34	13
$\Delta G n\text{-C}_6\text{H}_{14}(\text{g})^d$	0.80(±0.1)	-2.27(±0.6)	0.12(±0.10) <sup>e</sup>		0.61(±0.6) <sup>e</sup>	14	0.95	75	13
$\Delta G c\text{-C}_6\text{H}_{12}(\text{g})^d$	0.67(±0.1)	-1.87(±0.5)	0.16(±0.08) <sup>e</sup>		-0.02(±0.4) <sup>e</sup>	11	0.97	97	13
$\Delta G n\text{-C}_7\text{H}_{16}(\text{g})^d$	0.75(±0.2)	-2.11(±0.9)	0.25(±0.1) <sup>e</sup>		-0.21(±0.8) <sup>e</sup>	9	0.97	33	13
$\Delta G (\text{CH}_3)_4\text{Sn}(\text{g})^d$	0.65(±0.1)	-1.66(±0.4)	0.07(±0.1) <sup>e</sup>		0.01(±0.5) <sup>e</sup>	15	0.97	135	13
$\Delta G \text{ Et}_4\text{N}^+, \text{I}^-^c$	8.92(±1.7)	-58.59(±8.2)	-0.6(±0.8) <sup>e</sup>		99.59(±9.8)	12	0.98	149	27

<sup>a</sup> Six heats were solved for five unknowns. Reasonable values for the parameters constitutes the main check on this fit. CH<sub>2</sub>Cl<sub>2</sub> does not fit using a nonspecific model indicating  $\sim 7 \text{ kcal mol}^{-1}$  hydrogen bonding. <sup>b</sup> Loading of bovine serum albumin into microspheres of poly(D,L-lactide) from various solvents as the percent in the polymer. <sup>c</sup> Free energies of solution of the separated ion pair. <sup>d</sup> Free energies of solution of the gaseous solute at 25 °C. <sup>e</sup> This contribution is not statistically significant. <sup>f</sup> All  $F$  values are significant at the 99.5% confidence level.

the reported<sup>13</sup> free energies of solution of the nonpolar gaseous hydrocarbon solutes that have been studied in nine or more nonacidic solvents are given in Table 5. As expected for poorly basic solutes, the solubilities in CCl<sub>4</sub> solvent are predicted. The signs and trends in the solute parameters are those expected on the basis of the USCM.

The  $W$  for free energies of solution of cyclohexane can be compared to the  $W$  values for the enthalpies. Though the free energy of solution of gaseous C<sub>6</sub>H<sub>12</sub> in the  $S'$ ,  $E_B' = 0$  solvent is close to zero, the enthalpy is exothermic. Thus, the entropy of solution of this gas, like the entropy of condensation, is a

negative number whose  $T\Delta S$  value is close to  $\Delta H$  but opposite in sign. With a less exothermic contribution from the dispersion contribution to solvation, the free energies of solution of the smaller alkanes are dominated by the positive  $W$  resulting from the entropy term.

Polymers are an important class of solutes because solvent selection is important in polymer processing. Although the number of solvents studied is small, the model is applied to the enthalpies of solution of nitrocellulose<sup>24</sup> (Table 3) and the biodegradable poly(D,L-lactide) (PLA).<sup>28</sup> The limited data for the heat of solution can be fit to eq 6, Table 5. The signs on

the parameters are again as expected, and for PLA they indicate a large specific interaction of the solute with basic solvents that is largely electrostatic in nature.

The loading efficiency (LE) of BSA is the percent of the hydrophilic solute bovine serum albumin entrapped in microspheres of PLA in basic solvents.<sup>28</sup> Though the number of solvents is limited and the conclusions tentative, the problem is discussed here to illustrate the insights possible from the cavity USM analysis. The log LE can be predicted with eq 6. Positive signs (Table 5) increase loading efficiency, and negative signs lower it. Loading efficiency is a complex property, involving three competitive interactions. The interaction of BSA with the polymer, PLA, is a constant and should lead to a positive intercept. The interactions of both BSA with solvent and PLA with solvent should disfavor the loading (LE) of BSA. The meaning of the parameters from the fit of log LE are the net of these competitive interactions.

All of the BSA interactions with PLA are constant as the solvent is varied and incorporated into  $W$ . Thus, in addition to the usual  $W$  contribution there are also  $QS'^2_{\text{PLA}}$ ,  $P_{\text{BSA}}S'_{\text{PLA}}$ ,  $E_{\text{BSA}}E_{\text{PLA}}$ , and  $C_{\text{BSA}}C_{\text{PLA}}$  components. The solvent dependent component of LE is being fit to

$$\log \text{LE} = QS'^2_{\text{sol}} + PS'_{\text{sol}} + (E_{\text{BSA}} + E_{\text{PLA}})E_{\text{B}}^{\text{sol}} + (C_{\text{BSA}} + C_{\text{PLA}})C_{\text{B}}^{\text{sol}} + W \quad (7)$$

The signs of the parameters  $Q$ ,  $P$ ,  $(E_{\text{BSA}} + E_{\text{PLA}})$ , and  $(C_{\text{BSA}} + C_{\text{PLA}})$  are those expected for the solvent contribution. An increase of  $S'$  from 1.5 to 3.0 changes the net of  $0.29S'^2 - 2.09S'$  from  $-2.5$  to  $-3.66$  decreasing the LE. Thus, weakly basic, poorly solvating solvents should be employed for maximum loading efficiency.

## Conclusion

Solvent-dependent, physicochemical properties which have contributions from a solid, liquid, or gaseous solute dissolving in the solvent



can be treated with previously reported USM parameters by adding a term for creation of a cavity in the solvent. The solvent-dependent component of cavity formation is given by the square of  $S'$ , the solvent nonspecific solvation parameter. The solute dependent component is given by  $Q$ . The resulting unified solvation cavity model, USCM, is given by eq 6:

$$\Delta\chi = QS'^2 + PS' + E_{\text{A}}*E_{\text{B}} + C_{\text{A}}*C_{\text{B}} + W \quad (6)$$

where  $QS'^2$  is the energy to create the cavity,  $PS'$  is the nonspecific solvation of the solute by the solvent, and  $W$  is the value of  $\Delta\chi$  for the physicochemical property in a solvent with  $E_{\text{B}} = C_{\text{B}} = S' = 0$ . The energy to disperse a liquid or solid solute in a solvent is included in  $W$ . Not only does  $S'$  fit the previously reported spectral shifts for over 50 different solutes in over 30 different solvents, but now free energies and enthalpies of solution for 35 different solutes in a variety of solvents are correlated.

Any notion that this report is a fortuitous correlation of data, by having enough variables, is dispelled by the consistent pattern of endoergic cavity formation parameters,  $Q$ , endoergic liquid or solid solute dissociation parameters,  $W$ , and exothermic solvation parameters,  $P$ , resulting from the 35 different correlations attempted.

The excellent correlations provide confidence in the prediction of solubilities of solutes in solvents. The extensive effort that has gone into this most important problem is illustrated by the extensive literature on solvent polarity measurements and regular solution theory. These areas are combined into a single treatment by the results of this study. With the determination of solubilities of a new polymer or drug in a minimal number of solvents, the solubilities in all of the solvents whose  $S'$  values are known can be predicted confidently.

The model can be applied to complex chemical processes by determining all the energy components and treating each independently or by combining quantities with similar functional form into one parameter that gives the sum or difference in the quantity as shown in eq 7.

In addition to these important practical applications, patterns in the solute parameters from data fits have and, as new data become available, will continue to increase our understanding of the solution process. The relative importance of specific and nonspecific interactions in determining solubility can be determined. Exceptions to the correlations can, as was the case for MIm, show solute aggregation or subtle specific interactions. A very provocative suggestion is made by comparing the solute parameters from the USCM analysis of the systems in this study referred to as the dynamic cavity model. In this view, the energetics of cavity formation corresponds to the disruption of solvent order and aggregates that are influenced more by the polarity of the solute than by solute size.

**Acknowledgment.** P.B. thanks the support of Consejería de Educacion y Cultura of CAM and Comision Interministerial de Ciencia y Tecnologia (CICYT, project no. SAF94-1018, Spain). The authors thank Nam-Tran Ho for making experimental data on polylactide available before publication.

## References and Notes

- (1) Hidebrand, J. H.; Scott, R. L. *The Solubility of Nonelectrolytes*, 3rd ed.; Reinhold: New York, 1950.
- (2) Hildebrand, J. M.; Prausnitz, J. M.; Scott, R. L. *Regular and Related Solutions*; Van Nostrand Reinhold: New York, 1970; pp 21–22.
- (3) Chuu, M. S.; Knauss, C. J.; Ruch, R. J.; Myers, R. R. *ACS Symp. Ser.* **1983**, 227, 141 and references therein.
- (4) Fowkes, F. M.; Tischler, D. O.; Wolfe, J. A.; Lannigan, L. A.; Ademu-John, C. M.; Halliwell, M. J. *J. Polym. Sci., Polym. Chem. Ed.* **1984**, 106, 6638 and references therein.
- (5) (a) Bustamante, P.; Escalera, B.; Martin, A.; Selles, E. *J. Pharm. Sci.* **1989**, 78, 567. (b) Yang, S. S.; Guillory, J. K. *J. Pharm. Sci.* **1972**, 61, 26. (c) Beerbower, A.; Wu, P. L.; Martin, A. *J. Pharm. Sci.* **1984**, 73, 179. (d) Martin, A.; Wu, P. L.; Beerbower, A. *J. Pharm. Sci.* **1984**, 73, 199.
- (6) (a) Hansen, C.; Beerbower, A. In *Encyclopedia of Chemical Technology*; Solubility Parameters; Suppl. 2nd ed.; Wiley: New York, 1971. (b) Brandrup, J.; Immergut, E. H. *Polymer Handbook*, 3rd ed.; Wiley: New York, 1989.
- (7) Reiss, H. L. *Adv. Chem. Phys.* **1966**, 9, 1.
- (8) Leborvitz, J. L.; Hefland, E.; Proestgaard, J. *Chem. Phys.* **1965**, 43, 774.
- (9) Tully-Smith, D. M.; Reiss, H. *J. Chem. Phys.* **1971**, 55, 1674.
- (10) Pierotti, R. A. *Chem. Rev. (Washington, D.C.)* **1976**, 76, 717.
- (11) (a) Reichardt, C. *Solvents and Solvent Effects in Organic Chemistry*, 2nd ed.; VCH Publishers: Orlando, FL, 1988. (b) Reichardt, C. *Angew. Chem., Int. Ed. Engl.* **1965**, 4, 29.
- (12) (a) Herbrandson, H. F.; Neufeld, F. R. *J. Org. Chem.* **1966**, 31, 1140. (b) Kamlet, M. J.; Abboud, J. L.; Abraham, M. H.; Taft, R. W. *J. Org. Chem.* **1983**, 48, 2877 and references therein.
- (13) (a) Abraham, M. H. *J. Am. Chem. Soc.* **1982**, 104, 2085 and references therein. (b) Abraham, M. H. *J. Am. Chem. Soc.* **1979**, 101, 5477. (c) Kamlet, M. J.; Carr, P. W.; Taft, R. W.; Abraham, M. H. *J. Am. Chem. Soc.* **1981**, 103, 6062.
- (14) Drago, R. S. *Applications of Electrostatic-Covalent Models in Chemistry*; Surfside Scientific: Gainesville, FL, 1995.
- (15) (a) Fowkes, F. M. *Org. Coat. Plast. Chem.* **1979**, 40, 13. (b) Fowkes, F. M. In *Acid-Base Interactions*; Miteal, K. L., Anderson, H. R., Jr. Eds.; VSP: Utrecht, 1991.
- (16) Ho, N.-T. *J. Phys. Chem.* **1994**, 98, 5362.

- (17) Bustamante, P.; Jimenez-Duran, M.; Escalera, B. *Int. J. Pharm.* **1996**, 129, 283.
- (18) (a) Drago, R. S. *J. Chem. Soc., Perkin Trans. 2* **1992**, 1827. (b) Drago, R. S. *J. Org. Chem.* **1992**, 57, 6547. (c) Drago, R. S.; Hirsch, M. S.; Ferris, D. C.; Chronister, C. W. *J. Chem. Soc., Perkin Trans. 2* **1994**, 219. (d) Drago, R. S.; Kovala-Demertzi, D.; Ferris, D. C. *J. Coord. Chem.* **1994**, 32, 145. (e) Ferris, D. C.; Drago, R. S. *J. Am. Chem. Soc.* **1994**, 116, 7509. (f) Drago, R. S.; Ferris, D. C. *J. Phys. Chem.* **1995**, 99, 6563.
- (19) Substituting the value of  $\delta = 5.9$  for perfluorohexane into eq 4 leads to an  $S'$  value of  $-0.07$ . This should not be interpreted as a repulsion but as a problem arising from the lack of good data to clearly establish  $S' = 0$  in the reported unified solvation data fit. The true  $S' = 0$  of an adjusted scale would be even less solvating than perfluorohexane but more solvating than the gas phase.
- (20) (a) Spencer, J. N.; Berger, S. K.; Powell, C. R.; Henning, B. D.; Furman, G. S.; Loffredo, W. M.; Rydberg, E. M.; Neubert, R. A.; Shopp, G. E.; Blauch, D. N. *J. Phys. Chem.* **1981**, 85, 1236. (b) Spencer, J. N.; Gleim, J. E.; Blevins, C. H.; Garrett, R. C.; Mayer, F. J. *J. Phys. Chem.* **1979**, 83, 1249.
- (21) Krishnan, C. V.; Friedman, H. L. *J. Phys. Chem.* **1971**, 75, 3598.
- (22) (a) Catalán, J.; Gomez, J.; Saiz, J. L.; Couto, A.; Ferraris, M.; Laynez, J. *J. Chem. Soc., Perkin Trans. 2*, in press. (b) Catalán, J.; Gomez, J.; Couto, A.; Laynez, J. *J. Am. Chem. Soc.* **1990**, 112, 1678.
- (23) (a) Drago, R. S.; Nozari, M. S.; Vogel, G. C. *J. Am. Chem. Soc.* **1972**, 94, 90. (b) Nozari, M. S.; Drago, R. S. *J. Am. Chem. Soc.* **1972**, 94, 6877. (c) Nozari, M. S.; Jensen, C. D.; Drago, R. S. *J. Am. Chem. Soc.* **1973**, 95, 3162.
- (24) Slejko, F. L. Ph.D. Thesis, University of Illinois, 1972.
- (25) Doan, P. E.; Drago, R. S. *J. Am. Chem. Soc.* **1984**, 106, 2772.
- (26) Ideal solubility  $= \ln X_2^i = (-\Delta H^F/R)((1/T) - (1/T^F))$ , BA =  $-1.4908$ , PB =  $-4.8968$ , MB =  $-2.3146$ ,  $T$  = experimental  $T(298\text{ K})$ ,  $\Delta H^F$  = cal/mol, correlation of intercepts vs  $\ln X_2^i = 0.999$ .
- (27) (a) Abraham, M. H.; Grellier, P. L.; Nasehzadeh, A.; Walker, A. C. *J. Chem. Soc., Perkin Trans. 2* **1988**, 1717. (b) Abraham, M. H. *J. Chem. Soc., Perkin Trans. 2* **1972**, 1343.
- (28) Gander, B.; Merkle, H. P.; Nguyễn, V. P.; Hô, N.-T. *J. Phys. Chem.* **1995**, 99, 16144.
- (29) Arnett, E. M.; Joris, L.; Mitchell, E.; Murty, T. S. S. R.; Gorie, T. M.; Schleyer, P. V. R. *J. Am. Chem. Soc.* **1970**, 92, 2365 and references therein.