

UPPER III: Unified Physical Property Estimation Relationships. Application to Non-Hydrogen Bonding Aromatic Compounds

NEERA JAIN AND SAMUEL H. YALKOWSKY*

Contribution from *Department of Pharmaceutical Sciences, College of Pharmacy, 1703 E. Mabel Street, The University of Arizona, Tucson, Arizona 85721.*

Received April 9, 1999. Accepted for publication June 28, 1999.

Abstract □ The UPPER scheme uses four additive and two nonadditive parameters and several well-known equations to calculate 21 physical properties of organic compounds strictly from molecular structure. The scheme allows reasonable estimations of melting and boiling points, aqueous and octanol solubilities, air–octanol, air–water, and octanol–water partition coefficients, vapor pressure, and other properties. In this report non-hydrogen bonding aromatic compounds are used to evaluate a portion of the UPPER scheme.

Introduction

A number of group contribution methods of calculating various physical properties of compounds such as melting point, boiling point, aqueous solubility, octanol solubility, partition coefficients, and vapor pressure have been reported. Some of these have been reviewed by Lyman et al.¹ The most commonly used schemes for each property are independent of one another and are based on different models and assumptions for the molecule. Two notable exceptions have been presented by Bondi² and Joback and Reid.³ These each use a single molecular fragmentation scheme for the calculation of several physical properties. Recently computational approaches have been used by several groups.^{4–12} Unlike most existing schemes, the UPPER (unified physical property estimation relationships) scheme^{13,14} uses calculations that are simple, straightforward, and can be performed without a computer. They are also based upon well-known physicochemical relationships. The molecular descriptors used in the UPPER scheme are well defined and highly intuitive. The UPPER scheme is unique in its use of a combination of additive group contribution values and simple nonadditive molecular parameters that account for the effects of the overall molecular geometry.

The UPPER scheme has many applications in the fields of pharmaceutical, environmental, and material sciences. Estimations of the above physical properties can provide important insight regarding the pharmaceutical efficacy, environmental fate, and industrial utility of a compound.

In this study estimated values for eight physical properties calculated by UPPER are compared to experimental values obtained from the literature. Aromatic compounds (benzenes, naphthalenes, anthracenes, phenanthrenes, and biphenyls) substituted with non-hydrogen bonding groups

(i.e., methyl, fluoro, chloro, bromo, iodo, nitro) are used for this evaluation.

Theoretical Background

UPPER is used to calculate a number of physical properties directly from molecular structure. Four sets of group contribution values are used to calculate four additive physical properties: heat of boiling, heat of melting, molar volume, and the aqueous activity coefficient. The breakdown of the molecule into its constitutive groups is both simple and uniform for all the calculations. Additionally, two nonadditive molecular descriptors that account for molecular symmetry and flexibility are used to estimate transition entropies. These six parameters are then used to calculate several fundamental physical properties via well-accepted equations.

Additive Molecular Parameters—the change in enthalpy that accompanies a phase transition is the result of the total intermolecular interactions of the molecule. It can be assumed that a group (for example an aromatic CH₃) will make a nearly constant contribution to the molar enthalpy change. On this basis the enthalpies of boiling and melting can each be considered as the summation of group contributions from constituent groups. Thus the contribution of each group toward the molar heat of boiling and melting can be quantitated as the sum of b_i and m_i values, respectively. Simamora et al.¹⁵ and Krzyzaniak et al.¹⁶ used multiple linear regression to generate b_i and m_i values from literature data for thousands of aliphatic and aromatic compounds. The b_i and m_i values¹⁵ used in this study are shown in Table 1.

The molar volume of a liquid is the ratio of the molecular weight to the density. Each group occupies a nearly constant volume in a molecule that is designated as its v_i value. A number of group contribution schemes have been used to estimate the molar volume. In this study the values developed by Fedors¹⁷ are used and are shown in Table 1.

Myrdal et al.^{18–20} presented the AQUAFAC (aqueous functional activity coefficients) scheme that considers the group aqueous activity coefficient as an additive property. The AQUAFAC model is based on an extensive database (AQUASOL) of reported solubilities. The q_i values^{18–22} that are given in Table 1 are group contributions to the aqueous activity coefficient of the solute.

Nonadditive Molecular Parameters—The two non-additive molecular descriptors that account for the geometry of the molecule are the symmetry and flexibility numbers. Dannenfelser et al.²³ and Simamora et al.²⁴ have

* To whom correspondence should be addressed. Tel: (520) 626-1289, fax: (520) 626-4063, e-mail: yalkowsk@pharmacy.arizona.edu.

Table 1—Additive Parameters

group	description	b_i	m_i	v_i	q_i
CAR	substituted aromatic carbon	-704	97	-3.9	0.525
CHAR	unsubstituted aromatic carbon	5670	1940	15.1	0.321
CBR	CAR involved in the bridging of two rings	-704	97	-1.4	0.525
CBIP	CAR involved in a biphenyl linkage	-4890	-2140	-4.1	0.525
YCH3	methyl group attached to an aromatic ring	8040	2600	33.5	0.204
YF	fluorine attached to an aromatic ring	5890	1950	18.0	-0.141
YCl	chlorine attached to an aromatic ring	9330	3400	28.0	0.409
YBr	bromine attached to an aromatic ring	10960	3900	34.0	0.645
YI	iodine attached to an aromatic ring	13630	4440	35.5	0.887
YNO2	nitro group attached to an aromatic ring	13110	5072 ^a	32.0	0.082
OBIP	correction for ortho substituents in a biphenyl	-1630	-1190	0.0	-0.123

^a Note: This value is a misprint in ref 15.

Table 2—Molecular Symmetry Numbers (σ) of Some Compounds

σ	Examples
1	
2	
2	
3	
4	
6	
12	

described these parameters in detail. The external rotational symmetry number (σ) is a measure of the rotational degeneracy of the molecule. It is equal to the number of orientations of a molecule that are identical with some reference orientation. The assignment of molecular symmetry number is illustrated in Table 2. The flexibility (ϕ) of a molecule is a measure of the number of stable torsional conformations that it can assume. Since the molecules of this study are considered rigid, they are all assigned a flexibility number of unity. Note, biphenyls can assume

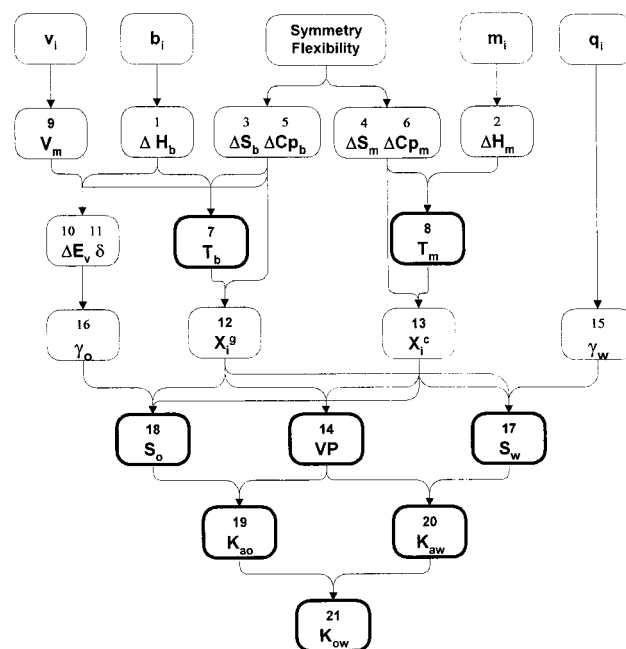


Figure 1—Schematic representation of UPPER.

more than one conformation; therefore, the entropy of fusion is approximated using the most stable conformation. The ortho substituents on the biphenyl can result in noncoplanarity of the molecule and therefore a correction factor, OBIP, is used as shown in Table 1.

Calculation of the Physical Properties—The logic of the UPPER scheme is illustrated in Figure 1. Twenty-one physical properties are each calculated from other properties by the use of simple, well-established relationships described below. The eight properties considered in this study are outlined in bold in Figure 1. The numbers in the following brief summary of the UPPER scheme correspond to the numbers in Figure 1.

1. *Heat of Boiling* (ΔH_b)—The heat of boiling, or the enthalpy of vaporization at the normal boiling temperature (J/mol), is calculated as the summation of the b_i values for the constitutive groups of the molecule. Therefore,

$$\Delta H_b = \sum (n_i b_i) \quad (1)$$

where n_i is the number of times the group b_i appears in the molecule.

2. *Heat of Melting* (ΔH_m)—The heat of melting, or the enthalpy of fusion (J/mol) is analogously calculated as the summation of the m_i values for the constitutive groups of the molecule, so that

$$\Delta H_m = \sum (n_i m_i) \quad (2)$$

3. *Entropy of Boiling* (ΔS_b)—The entropy of boiling, or the entropy of vaporization at the normal boiling point (J/K-mol) is calculated by Trouton's rule, which is applicable to coal tar derivatives, including non-hydrogen bonding aromatic compounds. Trouton's rule states that

$$\Delta S_b = 88 \quad (3)$$

4. *Entropy of Melting* (ΔS_m)—The entropy of melting, or the entropy of fusion (J/K-mol) is calculated by the Dannenfelser and Yalkowsky equation²³ that accounts for the effects of symmetry (σ) and flexibility (ϕ) upon the probability of a molecule being properly oriented and conformed for incorporation into the crystal lattice. The equation states:

$$\Delta S_m = 56.5 - 19.1 \log \sigma + 19.1 \log \phi \quad (4)$$

For the rigid molecules of this study $\phi = 1$ and eq 4 reduces to

$$\Delta S_m = 56.5 - 19.1 \log \sigma \quad (4a)$$

(Note that for most organic compounds the value of σ is either 1 or 2, and only 2% of compounds have σ exceeding 4.)

5. *Heat Capacity Changes on Boiling* (ΔC_{pb})—Myrdal et al.²⁵ estimated that the heat capacity change on boiling (J/K-mol) is related to the flexibility of the molecule as:

$$\Delta C_{pb} = -90 - 6.8 \log \phi \quad (5)$$

For the rigid molecules of this study $\phi = 1$, and the above eq 5 reduces to

$$\Delta C_{pb} = -90 \quad (5a)$$

6. *Heat Capacity Changes on Melting* (ΔC_{pm})—Mackay²⁶ and Yalkowsky and Mishra²⁷ have shown that the heat capacity change on melting can be approximated to be zero for rigid molecules, i.e.,

$$\Delta C_{pm} = 0 \quad (6)$$

7. *Boiling Point* (T_b)—The boiling point (K) is calculated as the ratio of the enthalpy of boiling (from eq 1) to the entropy of boiling (from eq 3). Thus

$$T_b = \frac{\Delta H_b}{\Delta S_b} = \frac{\sum n_i b_i}{88} \quad (7)$$

8. *Melting Point* (T_m)—The melting point (K) is calculated as the ratio of the enthalpy of melting (from eq 2) to the entropy of melting (from eq 4a).

$$T_m = \frac{\Delta H_m}{\Delta S_m} = \frac{\sum n_i m_i}{56.5 - 19.1 \log \sigma} \quad (8)$$

9. *Molar Volume* (V_m)—The molar volume (cm³/mol) is calculated as the summation of the v_i values.

$$V = \sum n_i v_i \quad (9)$$

10. *Energy of Vaporization* (ΔE_v)—According to the second law of thermodynamics

$$\Delta E_b = \Delta H_b - RT \quad (10)$$

where ΔE_b is the change in energy of boiling at the boiling point. The energy of vaporization, ΔE_v (J/mol) at any temperature is related to its value at the boiling point by

$$\Delta E_v = \Delta E_b - (T_b - T)\Delta C_{pb} \quad (10a)$$

where ΔH_b , ΔC_{pb} , and T_b are obtained from eqs 1, 5a, and 7, respectively, R is the gas constant, and T is 298 K. Combining eqs 10 and 10a gives

$$\Delta E_v = \Delta H_b - (T_b - T)\Delta C_{pb} - RT \quad (10b)$$

11. *Solubility Parameter* (δ)—The solubility parameter (J/cm³)^{0.5} is an expression of the cohesion between like molecules. It is calculated from the energy of vaporization (from eq 10b) and the molar volume (from eq 9) by

$$\delta = \sqrt{\frac{\Delta E_v}{V}} \quad (11)$$

12. *Ideal Solubility of a Gas* (X_i^g)—The ideal mole fractional solubility of a gas is the solubility that it would have in a perfect solvent. It is calculated using the integrated form of the Clausius–Clayperon equation as

$$\log X_i^g = \frac{\Delta S_b(T_b - T)}{2.3RT} - \frac{\Delta C_{pb}(T_b - T)}{2.3RT} + \frac{\Delta C_{pb}}{2.3R} \ln \frac{T_b}{T} \quad (12)$$

where ΔS_b , C_{pb} , and T_b are calculated from eqs 3, 5a, and 7, respectively.

13. *Ideal Solubility of a Crystal* (X_i^c)—The ideal mole fractional solubility of a crystalline solid is similarly calculated from the van't Hoff equation

$$\log X_i^c = -\frac{\Delta S_m(T_m - T)}{2.3RT} - \frac{\Delta C_{pm}(T_m - T)}{2.3RT} + \frac{\Delta C_{pm}}{2.3R} \ln \frac{T_m}{T} \quad (13)$$

According to eq 6, $\Delta C_{pm} = 0$; therefore, eq 13 can be simplified to

$$\log X_i^c = -\frac{\Delta S_m(T_m - T)}{2.3RT} \quad (13a)$$

where ΔS_m and T_m are calculated from eqs 4a and 8, respectively. Note that when the melting point of a compound is less than 298 K, the logarithm of the ideal solubility of the crystal is zero or in other words it is completely miscible. Also note that for solids of low symmetry the above eq 13a is approximately

$$\log X_i^c = -0.01(T_m - T) \quad (13b)$$

14. *Vapor Pressure* (VP)—Yalkowsky and Mishra²⁸ showed that the vapor pressure (atm) can be calculated from the integrated form of the Clausius–Clapeyron equation

$$\log VP = -\frac{\Delta S_m(T_m - T)}{2.3RT} - \frac{\Delta C_{pm}(T_m - T)}{2.3RT} + \frac{\Delta C_{pm}}{2.3RT} \ln \frac{T_m}{T} - \frac{\Delta S_b(T_b - T)}{2.3RT} + \frac{\Delta C_{pb}(T_m - T)}{2.3RT} - \frac{\Delta C_{pb}}{2.3RT} \ln \frac{T_b}{T} \quad (14)$$

By incorporating eqs 12 and 13 the above equation can be written as

$$\log VP = \log X_i^c - \log X_i^g \quad (14a)$$

If the compound is a liquid, eq 14a is reduced to eq 14b as the crystal term drops out.

$$\log VP = -\log X_i^g \quad (14b)$$

15. *Activity Coefficient in Water (γ_w)*—The logarithm of the aqueous activity coefficient (mol/L) is equal to the summation of the AQUAFAC substituent group activity coefficients.

$$\log \gamma_w = \sum n_i q_i \quad (15)$$

16. *Activity Coefficient in Octanol (γ_o)*—The octanol activity coefficient (mol/L) is calculated by applying the Scatchard–Hildebrand relationship to octanol. This gives

$$\log \gamma_o = \frac{V_u(\delta_{\text{oct}} - \delta_u)^2 \phi_{\text{oct}}^2}{2.3RT} \quad (16)$$

where V_u and δ_u are the molar volume and solubility parameter of the solute calculated using eqs 9 and 11, respectively, and ϕ_{oct} is the volume fraction of the solvent, octanol. Using the solubility parameter for octanol as the solvent, eq 16 becomes

$$\log \gamma_{\text{oct}} = \frac{V_u(21.1 - \delta_u)^2 \phi_{\text{oct}}^2}{2.3RT} \quad (16a)$$

Hildebrand and Scott²⁹ showed that the critical temperature, T_c , at which two liquids are completely miscible, is

$$T_c = \frac{V(\delta_v - \delta_u)^2}{2R} \quad (16b)$$

where V is the arithmetic of the molar volumes of octanol and the solute. If we assume the molar volume of the solute is nearly same as octanol, 138 cm³/mol, i.e.,

$$V = \frac{V_{\text{oct}} + V_u}{2} \approx 138 \quad (16c)$$

then complete miscibility will be achieved if

$$|21.1 - \delta_u| < \sqrt{\frac{2RT}{138}} \quad (16d)$$

which corresponds to $15.1 < \delta_u < 27.1$ in (J/cm³)^{0.5} at 298 K or higher temperature. As most organic compounds have solubility parameters in this range, they are completely miscible with octanol.

17. *Solubility in Water (S_w)*—The aqueous solubility (mol/L) for solids is calculated as the ratio of the ideal solubility of the solute to its aqueous activity coefficient. In logarithmic terms this is

$$\log S_w = \log X_i^c - \log \gamma_w \quad (17)$$

where $\log X_i^c$ and $\log \gamma_w$ are calculated from eqs 13a and 15, respectively. For liquids X_i^c is set equal to unity and thus eq 17 reduces to

$$\log S_w = -\log \gamma_w \quad (17a)$$

18. *Solubility in Octanol (S_o)*—Likewise the octanol solubility (mol/L) is calculated as the ratio of the ideal solubility of the solute to its octanol activity coefficient. For a solid solute this is

$$\log S_o = \log X_i^c - \log \gamma_o \quad (18)$$

where $\log X_i^c$, and $\log \gamma_o$ are calculated from eqs 13a and 16, respectively.

19. *Air–Octanol Partition Coefficient (K_{ao})*—The air–octanol partition coefficient (atm–L/mol) is calculated as the ratio of the vapor pressure to the octanol solubility.

$$\log K_{ao} = \log VP - \log S_o \quad (19)$$

where $\log VP$, and $\log S_o$ are calculated from eqs 14a and 18, respectively.

20. *Air–Water Partition Coefficient or Henry's Law Constant (K_{aw})*—The air–water partition coefficient (atm–L/mol) is calculated as the ratio vapor pressure coefficient to the aqueous solubility.

$$\log K_{aw} = \log VP - \log S_w \quad (20)$$

where $\log VP$ and $\log S_w$ are calculated from eqs 14a and 17, respectively.

21. *Octanol–Water Partition Coefficient (K_{ow})*—The octanol–water partition coefficient is calculated as the ratio of the aqueous activity coefficient to the octanol activity coefficient.

$$\log K_{ow} = \log \gamma_w - \log \gamma_o \quad (21)$$

where $\log \gamma_w$ and $\log \gamma_o$ are calculated from eqs 15 and 16, respectively. Alternatively $\log K_{ow}$ can be calculated either by

$$\log K_{ow} = \log S_o - \log S_w \quad (21a)$$

$$\log K_{ow} = \log K_{aw} - \log K_{ao} \quad (21b)$$

where $\log S_w$, $\log S_o$, $\log K_{ao}$, and $\log K_{aw}$ are calculated from eqs 17, 18, 19, and 20, respectively.

Data Collection

Four hundred and five rigid, non-hydrogen bonding aromatic compounds including benzenes, naphthalenes, anthracenes, phenanthrenes, and biphenyls substituted with methyl, fluoro, chloro, bromo, iodo, and nitro groups were used in this study. Data for melting and boiling points,^{30–33} aqueous,^{31,33–37} and octanol³⁸ solubilities, air–water,³³ air–octanol,⁴² and octanol–water^{33,43} partition coefficients, and vapor pressure^{31,33,39–41} were taken from the literature.

Additive Parameters—Simamora et al.,¹⁵ Fedors,¹⁷ Myrdal et al.,^{18,19} and Lee et al.²¹ reported the values of the additive parameters used in this study. These are shown in Table 1.

Nonadditive Parameters—Symmetry numbers as described by Dannenfelser et al.²³ and illustrated in Table 2 were assigned to each compound. Since the substituted aromatics considered are rigid molecules ϕ is equal to unity for all compounds.

Results and Discussion

Boiling Point—The observed and predicted boiling points are in good agreement as evidenced by the fit of the data to the line of identity in Figure 2. The average

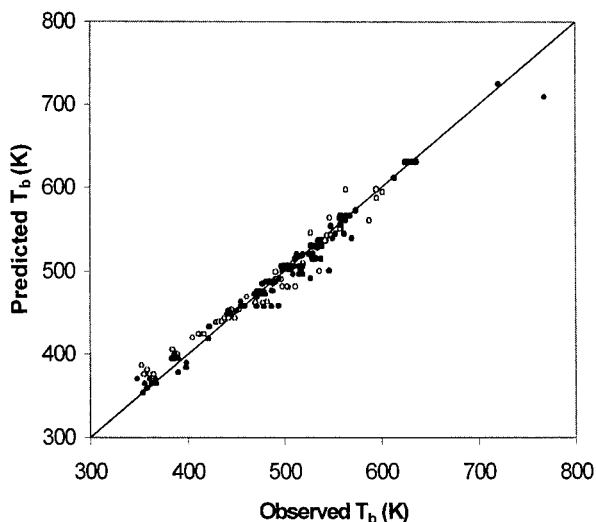


Figure 2—Observed vs predicted boiling point (K). (○) Compounds used in the training set to generate b_i values. (●) Compounds not used in the training set.

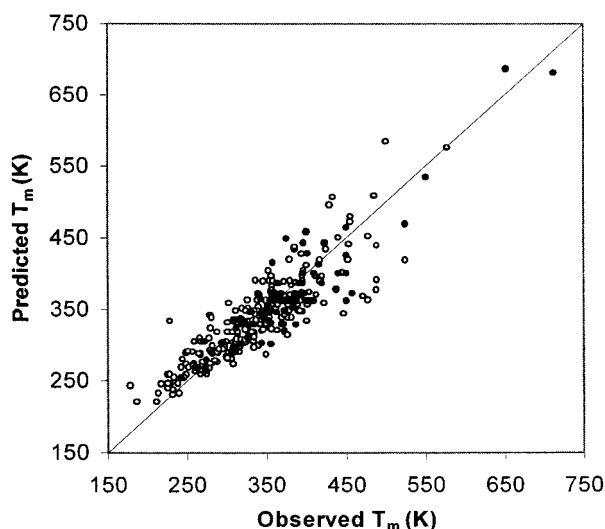


Figure 3—Observed vs predicted melting point (K). (○) Compounds used in the training set to generate m_i values. (●) compounds not used in the training set.

absolute error is 8.67 K for 191 compounds. Only 68 of the 191 compounds were used by Simamora et al.¹⁵ to generate the b_i values. These are denoted by open circles. The remainder of the compounds are a true test and are denoted by filled circles in Figure 2.

Melting Point—The observed and predicted melting points are shown in Figure 3. The average absolute error is 23 K for 338 compounds. Over 85 of these compounds were not used by Simamora et al.¹⁵ to generate m_i values. These compounds are represented by filled circles and are a true test set for the relationship.

As melting point is dependent upon the arrangement of the molecules in the crystal lattice as well as upon the strength of the pairwise group interactions, it is more difficult to predict than boiling point. The importance of incorporating a nonadditive parameter along with the group contributions is evident from the fact that the calculation of the melting points without a symmetry term yields a average absolute error of 30 K. The calculation of the melting point by group contributions alone leads to the same estimated melting point for constitutional isomers while using the symmetry number distinguishes them effectively.

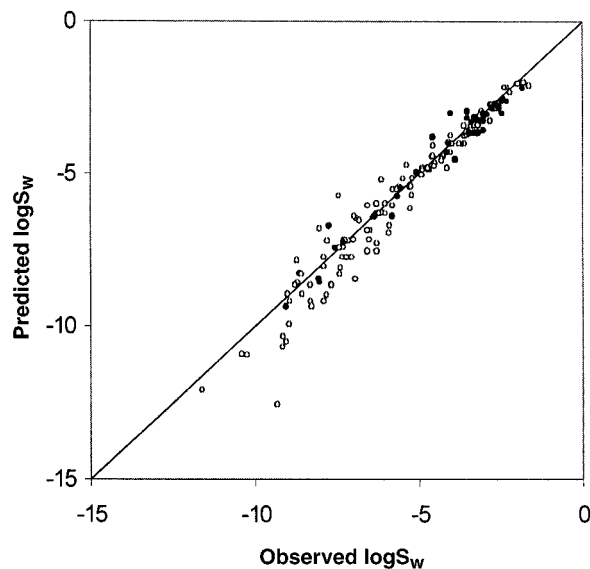


Figure 4—Observed vs predicted logarithm of molar aqueous solubility. (○) Compounds used in the training set to generate q_i values. (●) Compounds not used in the training set.

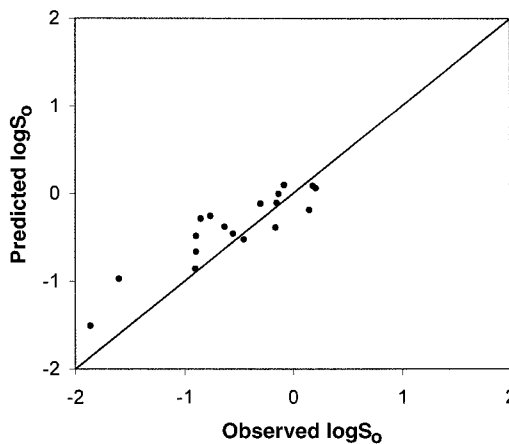


Figure 5—Observed vs predicted logarithm of molar octanol solubility.

Aqueous Solubility—The calculated solubility values agree very well with the experimental data as shown in Figure 4. The average absolute error of 0.38 log units for 165 compounds varying over 10 orders of magnitude is slightly less than a factor of 2.5. This is well within experimental error of the reported data. One hundred and thirty six compounds were used to generate q_i values by Myrdal et al.^{18,19} and Lee et al.²¹ using the true melting points. It should be noted that in this study the aqueous solubility was calculated using the calculated melting points. Therefore the filled circles in Figure 4 represent a true test set of melting point in determining solubility.

Octanol Solubility—Although not much data is available for octanol solubility, UPPER appears to be applicable to its prediction. The 22 reported octanol solubilities are in reasonable agreement with the calculated values as shown in Figure 5. The average absolute error is 0.40 log units (or a factor of 2.5) for the 22 compounds. Note that since no values were generated from octanol solubility data, the whole data set represents a true test for eq 18.

Vapor Pressure—Although no vapor pressure data has been used in the generation of the m_i or b_i coefficients in Table 1, the calculated vapor pressures are in very good agreement with the reported values as evident from Figure 6. Of the compounds shown, 27 were used by Myrdal et al.²⁵ to generate the value of ΔC_{p_i} , the heat capacity change

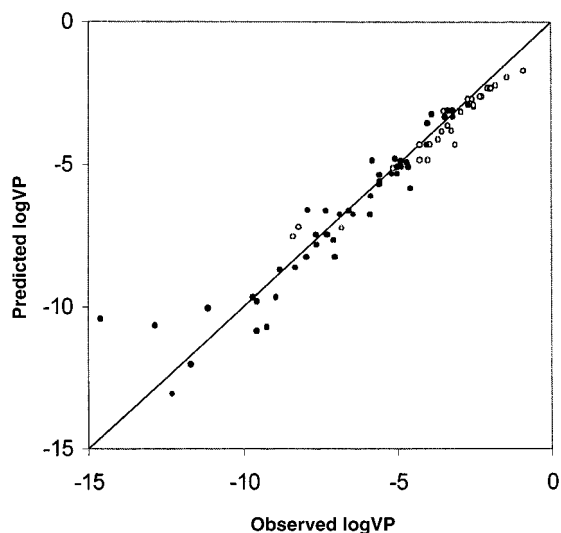


Figure 6—Observed vs predicted logarithm of vapor pressure (atm). (○) Compounds used in the training set to generate ΔC_{pb} ; (●) Compounds not used in the training set.

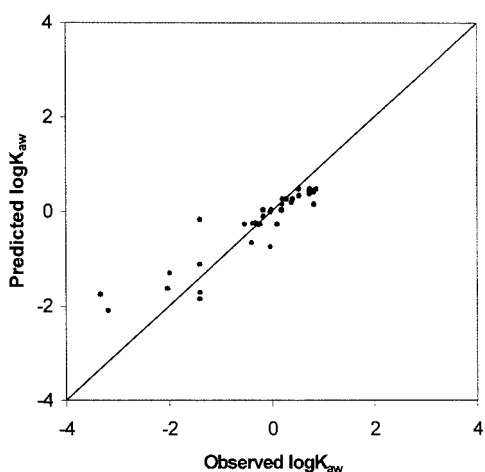


Figure 7—Observed vs predicted logarithm of Henry's Law Constant (atm-L/mol).

on boiling. The average absolute error is 0.38 log units for 73 compounds with vapor pressures covering 11 orders of magnitude. This is well within the error associated with vapor pressure measurements.

Air–Water Partition Coefficient or Henry's Law Constant—Most of the predicted values of the Henry's law constants for 37 compounds show good agreement with the experimentally determined values. However, the error for four compounds is unacceptably large. Because of these points the average absolute error for total test set is 0.49 log units. The data are plotted in Figure 7.

Air–Octanol Partition Coefficient—The calculated air–water partition coefficients of 25 compounds are in reasonable agreement with the experimental data reported⁴² with an average absolute error of 0.45 as shown in Figure 8.

Octanol–Water Partition Coefficient—The agreement between measured and estimated partition coefficients is shown in Figure 9. The average absolute error of 0.40 is primarily due to over estimation of partition coefficients greater than five. The good agreement between the observed and the predicted partition coefficient values below 10^5 (average absolute error of 0.25) is noteworthy since none of the parameters used for the estimations are based upon partitioning data. A note should be made that

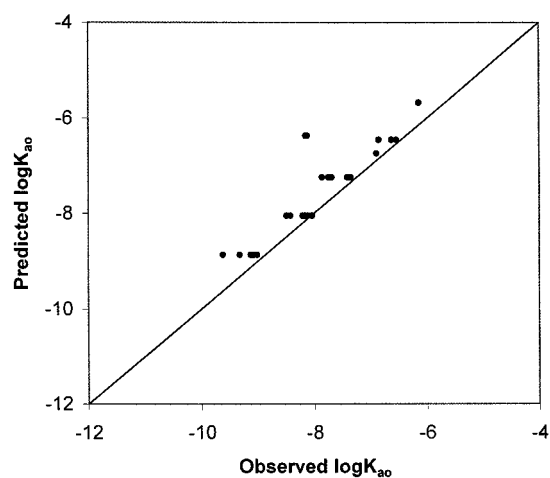


Figure 8—Observed vs predicted logarithm of air–octanol partition coefficient (atm-L/mol).

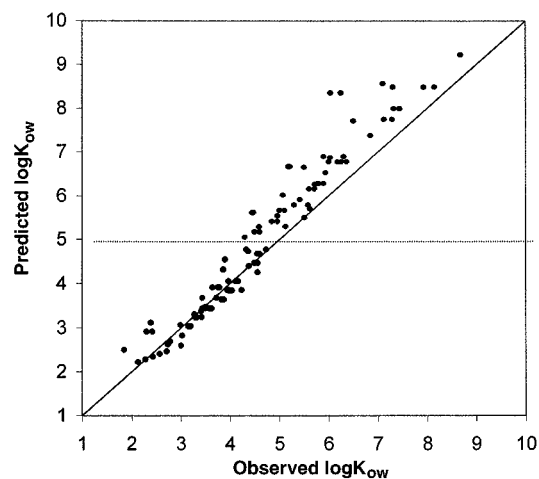


Figure 9—Observed vs predicted logarithm of the octanol–water partition coefficient.

Table 3—Summary of Physical Property Estimation Results

property	n	average absolute error	range	
			min	max
T_b	191	8.67	348.00	636.00
T_m	338	23.12	178.00	651.00
$\log S_w$	165	0.38	-11.62	-1.64
$\log S_o$	22	0.40	-2.77	0.21
$\log VP$	73	0.38	-12.30	-0.91
$\log K_{aw}$	37	0.49	-1.40	0.88
$\log K_{ao}$	25	0.42	-9.62	-6.14
$\log K_{ow}$	128	0.40	1.85	11.46
$\log K_{ow} (<10^5)$	84	0.25	1.85	5.00

the octanol–water partition coefficient calculated using eq 21 does not take into account the mutual solubilities of octanol and water. This is probably one of the reasons for the overestimations that are especially pronounced with high partition coefficients.

Conclusion

Table 3 summarizes the results of estimating the eight properties considered in this report. For each property it gives the number of compounds studied and the average absolute error of the estimate along with the range of the experimental values. The average absolute error of all the

Table 4—Observed and Predicted Properties for 66 Compounds

name	obsd T_m	pred T_m	obsd T_b	pred T_b	obsd log VP	pred log VP	obsd log S_w	pred log S_w	obsd log K_{ow}	pred log K_{ow}	obsd log K_{aw}	pred log K_{aw}	obsd log S_o	pred log S_o	obsd log K_{ao}	pred log K_{ao}
benzene	279	324	353	387	-0.91	-1.71	-1.64	-2.09	2.13	2.23	0.74	0.38			0.13	
toluenelene	178	244	384	406	-1.43	-1.91	-2.21	-2.33	2.73	2.63	0.83	0.42			0.30	
1,2-dimethylbenzene	248	259	417	424	-2.06	-2.29	-2.80	-2.74	3.15	3.04	0.75	0.46			0.30	
1,3-dimethylbenzene	225	259	412	424	-1.96	-2.29	-2.82	-2.74	3.20	3.04	0.85	0.46			0.30	
1,4-dimethylbenzene	286	292	411	424	-1.94	-2.29	-2.77	-2.74	3.18	3.04	0.76	0.46			0.30	
1,2,3,-trimethylbenzene	248	274	449	443	-2.70	-2.67	-3.20	-3.15	3.59	3.45	0.53	0.48			0.30	
1,2,4-trimethylbenzene	229	246	442	443	-2.57	-2.67	-3.31	-3.15	3.63	3.45	0.75	0.48			0.30	
1,3,5-trimethylbenzene	228	334	438	443	-2.50	-2.94	-3.40	-3.41	3.58	3.45	0.88	0.48			0.04	
1,2,4,5-tetramethylbenzene	353	326	470	462	-3.18	-3.29	-4.59	-3.78	4.00	3.86	1.40	0.49			0.08	
pentamethylbenzene	323	304	504	481	-4.03	-3.52	-4.00	-4.02	4.56	4.27		0.50			0.25	
chlorobenzene	228	260	405	420	-1.79	-2.20	-2.41	-2.54	3.02	2.84	0.54	0.34			0.30	
1,2-dichlorobenzene	256	291	453	454	-2.68	-2.89	-3.02	-3.15	3.44	3.45	0.22	0.26			0.30	
1,3-dichlorobenzene	249	291	445	454	-2.52	-2.89	-3.07	-3.15	3.49	3.45	0.41	0.26		0.21	0.30	
1,4-dichlorobenzene	327	328	446	454	-2.92	-3.12	-3.31	-3.39	3.44	3.45	0.30	0.26			0.06	
1,2,3,-trichlorobenzene	326	321	491	487	-3.42	-3.81	-4.10	-3.97	4.11	4.07	0.22	0.16			0.09	
1,2,4-trichlorobenzene	289	289	487	487	-3.36	-3.60	-3.61	-3.77	3.97	4.07	0.22	0.16			0.30	
pentachlorobenzene	357	383	548	555	-4.58	-5.85	-5.66	-5.75	5.12	5.29	-0.15	-0.10	-0.55		-0.45	
hexachlorobenzene	500	585	596	588	-7.07	-7.67	-7.56	-7.41	5.41	5.90	-0.24	-0.26	-1.86		-1.51	
fluorobenzene	231	231	358	381		-1.44	-1.80	-1.99	2.28	2.29		0.55			0.30	
1,2,3,5-tetrafluorobenzene		238	356	365		-1.13	-2.31	-2.18	2.71	2.48		1.05			0.30	
1,2,4,5-tetrafluorobenzene	277	268	363	365		-1.13	-2.38	-2.18	2.71	2.48		1.05			0.30	
bromobenzene	242	270	429	439	-2.26	-2.57	-2.55	-2.78	2.99	3.08	0.39	0.20			0.30	
1,2-dibromobenzene	277	310	497	491		-3.79	-3.50	-3.73	3.64	3.92		-0.05			0.19	
1,3-dibromobenzene	266	310	491	491	-3.24	-3.79	-3.54	-3.73	3.75	3.92		-0.05			0.19	
iodobenzene	244	281	461	469	-3.88	-3.21	-3.04	-3.02	3.27	3.32		-0.19			0.30	
1,4-diiodobenzene	404	374	558	552		-5.62	-5.37	-4.71	4.39	4.41			-0.91			
nitrobenzene	278	293	483	463		-3.08	-1.83	-2.21	1.85	2.51		-0.87			0.30	
1,2-dinitrobenzene	390	357		540		-5.28	-4.04	-3.02	1.69	2.80		-2.26			-0.22	
1,3-dinitrobenzene	363	357	570	540		-5.28	-2.46	-3.02	1.49	2.80		-2.26			-0.22	
1,4-dinitrobenzene	445	402		540		-5.58	-3.38	-3.32	1.47	2.80		-2.26			-0.52	
<i>o</i> -chlorotoluene	238	247	432	439	-2.30	-2.58	-3.52	-2.95	3.42	3.25		0.36			0.30	
<i>m</i> -chlorotoluene	225	247	435	439		-2.58	-3.52	-2.95	3.28	3.25		0.36			0.30	
<i>o</i> -fluorotoluene	211	221	386	400		-1.80		-2.40	2.78	2.70		0.60			0.30	
<i>m</i> -fluorotoluene	186	221	388	400		-1.80		-2.40	2.78	2.70		0.60			0.30	
<i>m</i> -bromotoluene	233	256	457	458		-2.97	-3.52	-3.18	3.50	3.48		0.22			0.30	
<i>o</i> -nitrotoluene	269	277	498	482		-3.49	-2.31	-2.62	2.30	2.92		-0.87			0.30	
<i>m</i> -nitrotoluene	288	277	503	482		-3.49	-2.46	-2.62	2.42	2.92		-0.87			0.30	
1-chloro-4-nitrobenzene	356	324	515	497		-4.03	-2.92	-3.05	2.39	3.13		-0.98			0.07	
2-bromochlorobenzene	261	270	477	472		-3.28	-3.19	-3.39	3.44	3.69		0.11			0.30	
3-bromochlorobenzene	252	270	469	472		-3.28	-3.21	-3.39	3.72	3.69		0.11			0.30	
1-fluoro-4-iodobenzene	300	283	455	464		-3.09	-3.13	-3.08	3.41	3.38		-0.01			0.30	
1-methylnaphthalene	251	292	513	518	-4.29	-4.28	-3.70	-4.03	3.87	4.33	-0.35	-0.25			0.30	
2-methylnaphthalene	307	292	514	518	-4.05	-4.28	-3.77	-4.03	3.86	4.33	-0.29	-0.25			0.30	
1,4-dimethylnaphthalene	281	339	535	537	-4.65	-5.07	-4.14	-4.80	4.37	4.73	-0.51	-0.27			-0.07	
anthracene	489	440	613	612	-8.40	-7.55	-6.39	-6.43	4.45	5.61	-1.40	-1.12	-1.91		-0.82	
9-methylanthracene	352	405		631	-7.63	-7.83	-5.89	-6.67	5.07	6.02		-1.16			-0.65	
phenanthrene	372	390	613	612	-8.40	-7.25	-5.26	-6.13	4.47	5.61	-1.49	-1.12	-0.45		-0.52	
biphenyl	344	336	527	531	-4.89	-4.86	-4.31	-4.56	3.90	4.56	-1.54	-0.30	-0.13		0.00	
<i>o</i> -chlorobiphenyl	307	274	527	546	-4.69	-4.90	-4.54	-4.75	4.30	5.05		-0.15			0.30	
<i>m</i> -chlorobiphenyl	290	295	547	565	-5.00	-5.32	-4.88	-4.87	4.60	5.17		-0.45			0.30	
<i>p</i> -chlorobiphenyl	348	329	557	565	-5.57	-5.60	-5.20	-5.15	4.50	5.17		-0.45			0.03	
4,4'-dichlorobiphenyl	422	359	564	598	-7.32	-6.65	-6.56	-6.03	5.30	5.79		-0.62			-0.25	
2,2'-dichlorobiphenyl	334	312	588	561	-5.58	-5.37	-5.27	-5.37	4.90	5.54		-0.01			0.17	
3,3'-dichlorobiphenyl	302	359		598	-6.57	-6.65	-5.80	-6.03	5.30	5.79		-0.62			-0.25	
2,2',3,3',4,4',6'-heptachlorobiphenyl	395	397		711	-9.56	-9.79	-8.30	-9.17	6.70	8.48		-0.62			-0.68	
2,2',3,3',5,5',6,6'-octachlorobiphenyl	434	507		726	-9.58	-10.83	-9.15	-10.32		8.97		-0.50			-1.35	
2,2',3,3',4,4',5,5',6-nonachlorobiphenyl	478	452		778	-11.71	-12.01	-10.26	-10.94		9.71		-1.07			-1.23	
2,2',3,3',4,4',5,5',6,6'-decachlorobiphenyl	578	576		793	-12.30	-13.06	-11.62	-12.10		10.20		-0.96	-2.77		-1.90	
<i>p</i> -bromobiphenyl		338		583		-6.11	-5.55	-5.47	4.96	5.41		-0.64			-0.06	
2,4,6-tribromobiphenyl	339	373		650		-8.00	-7.30	-7.23	6.03	6.86		-0.77			-0.36	
fluoranthene	383	354		596	-7.91	-6.61	-5.92	-6.91	5.22	6.66	-0.76	-0.25	-0.76		-0.25	-8.12
pyrene	423	444		596	-8.22	-7.21	-6.18	-7.51	5.18	6.66	-0.90	-0.85	-0.90		-0.85	-8.15
chrysene	525	470	721	725	-9.24	-10.69		-8.54		7.30	-2.60	-1.23	-2.60		-1.23	
perylene	551	535		709	-12.85	-10.63	-8.79	-9.92	6.25	8.35	-2.52	-1.57	-2.52		-1.57	
benzo[<i>a</i>]pyrene	451	426	768	709	-11.15	-10.03	-7.82	-9.32	6.04	8.35	-1.60	-0.97	-1.60		-0.97	
coronene	715	681			-14.60	-10.39	-9.33	-12.56		10.45	-2.37	-2.11	-2.37		-2.11	

logarithmic terms are generally of the same magnitude as the error associated with the experimental measurements. On the other hand the errors in estimation of the transition

temperatures exceed the accuracy of their measurement. However, these errors are less than those produced by any other predictive method and correspond to relatively small

errors in those logarithmic terms that are dependent upon the melting point. Therefore, since none of the average absolute errors exceeds half an order of magnitude, UPPER can be expected to consistently provide reasonable (order of magnitude) estimations of these properties.

The success of UPPER is due to three factors: its use of well-established physicochemical relationships, its use of a uniform breakdown of molecules into substituents, and its use of nonadditive parameters. Because all property values are calculated strictly from thermodynamically sound relationships, UPPER itself is thermodynamically sound. The use of mathematical relationships, which relate one parameter to another, minimizes the need for approximations and eliminates the need to utilize multiple structural breakdown schemes.

The use of nonadditive parameters to account for those molecular properties that are not simply the sum of group values is the most unique feature of UPPER. In the present study the use of molecular symmetry improves the estimation of melting point and enables the distinction of constitutional isomers. Consequently, it improves the estimation of solubility and vapor pressure which are both calculated from the melting point. Note that although a 30 K error in melting point corresponds to a 0.3 log unit error (i.e., a factor of 2) in either solubility or vapor pressure. An error in the estimation of the melting point has no effect on either air-liquid or liquid-liquid partition coefficients.

There were no new fitted parameters generated or used in this study. About one-third of the data set used in this study is a true test of the group contribution values previously determined by Simamora,¹⁵ Myrdal et al.,^{18,19} and Lee et al.²¹ A total of 119 boiling points, 86 melting points, 29 aqueous solubilities, and 40 vapor pressures were calculated strictly from the molecular structure as they were not used in generating the group contributions previously. Note that in this study the aqueous solubility and vapor pressure were calculated from the calculated melting and boiling points while both AQUAFAC coefficients and the heat capacity change on boiling were generated using the true melting and boiling temperatures. The excellent agreement of the aqueous solubility and vapor pressure values spanning almost 10 orders of magnitude is noteworthy and suggests that UPPER can be used to estimate values before actual costly experimentation. Even higher accuracy can be achieved with the use of the true melting point, an easily measurable property. For example, the use of the true melting point gives an absolute average error of 0.30.

The agreement of both the calculated air-water and octanol-water partition coefficients reflects the strength of the UPPER scheme as no partitioning data were used in the generation of the group contribution values of Table 1.

Thus the overall UPPER scheme does remarkably well as validated using 974 values for 8 biologically and environmentally relevant properties of 405 rigid, non-hydrogen bonding aromatic compounds. The logarithmic values in Table 3 have average absolute errors that correspond to less than a factor of 3 for all the properties and less than a factor of 2 for most of the properties.

A representative set of 66 compounds is presented in Table 4. All the data for 405 compounds are provided in Supporting Information.

Supporting Information Available—Table of observed and predicted properties. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

1. Lyman, W. J.; Reehl, W.; Rosenbalt, D. H. *Handbook of Chemical Property Estimations*; McGraw-Hill Book Co.: New York, 1982.
2. Bondi, A. *Physical Properties of Molecular Crystals, Liquids, and Glasses*; John Wiley and Sons: New York, 1968.
3. Joback, K. G. *A Unified Approach to Physical Property Estimation Using Multivariate Statistical Techniques*. Ph.D. Thesis. Massachusetts Institute of Technology, Boston, 1984.
4. Todeschini, R.; Gramatica, P.; Provenzani, R.; Marengo, E. Weighted holistic molecular descriptors. Part 2. Theory development and applications on modeling physicochemical properties of polyaromatic hydrocarbons. *Chemom. Intell. Lab. Syst.* **1995**, *27*, 221–229.
5. Mitchell, B. E.; Jurs, P. C. Prediction of Aqueous Solubility of Organic Compounds from Molecular Structure. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 489–496.
6. Mitchell, B. E.; Jurs, P. C. Prediction of Infinite Dilution Activity Coefficients of Organic Compounds in aqueous Solutions from Molecular Structure. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 200–209.
7. Katritzky, A. R.; Lobanov, V.S.; Karelson, M. Normal Boiling Points for Organic Compounds: Correlation and Prediction by a Quantitative Structure-Property Relationship. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 28–41.
8. Huibers, P. D. T.; Katritzky, A. R. Correlation of Hydrocarbons and Halogenated Hydrocarbons with Molecular Structure. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 283–292.
9. Katritzky, A. R.; Wang, Y.; Sild, S.; Tamm, t. QSPR Studies on Vapor Pressure, Aqueous Solubility, and the prediction of Water-Air Partition Coefficients. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 720–725.
10. Karickhoff, S. W.; Careirra, L. A.; Hilal, S. H. Prediction of Pollutant Physical Properties by Computer (SPARC). *Abstracts of Papers, Part 1*, 210th National Meeting of the American Chemical Society, Chicago, IL, Aug 20–24, 1995; American Chemical Society: Washington, DC, 1995; ENVR 19.
11. Hilal, S.H.; Carreira, L. A.; Karickhoff, S. W. Vapor pressure, Boiling Point and Activity Coefficient Calculation by SPARC. *Abstracts of Papers, Part 1*, 212th National Meeting of the American Chemical Society, Orlando, FL, Aug 25–29, 1996; American Chemical Society: Washington, DC, 1996; COMP 114.
12. Pearlman, R. S.; Smith, K. M. Novel algorithms for the design of diverse and focused combinatorial libraries. *Abstracts of Papers, Part 1*, 217th National Meeting of the American Chemical Society, Anaheim, GA, Mar 21–25, 1999; American Chemical Society: Washington, DC, 1999; U2698.
13. Yalkowsky, S. H.; Dannenfelser, R. M.; Myrdal, P.; Simamora, P. Unified Physical Property Estimation Relationships (UPPER). *Chemosphere* **1994**, *28*, 1657–1673.
14. Yalkowsky, S. H.; Myrdal, P.; Dannenfelser, R. M.; Simamora, P. UPPER II: Calculation of Physical Properties of the Chlorobenzenes. *Chemosphere* **1994**, *28*, 1675–1688.
15. Simamora, P.; Yalkowsky, S. H. Group Contribution Methods for Predicting the Melting Points and Boiling Points of Aromatic Compounds. *Ind. Eng. Chem. Res.* **1994**, *33*, 1405–1409.
16. Krzyzaniak, J. F.; Myrdal, P. B.; Simamora, P.; Yalkowsky, S. H. Boiling Point and Melting Point Prediction for Non-Hydrogen Bonding Compounds. *Ind. Eng. Chem. Res.* **1995**, *34*, 2531–2535.
17. Fedors, R. F. A Method for estimating Both the Solubility Parameters and Molar Volumes of Liquids. *Polym. Eng. Sci.* **1974**, *14*, 147–154.
18. Myrdal, P.; Ward, G. H.; Dannenfelser, R. M.; Mishra, D.; Yalkowsky, S. H. AQUAFAC 1: Aqueous Functional Group Activity Coefficients; Application to Hydrocarbons. *Chemosphere* **1992**, *24*, 1047–1061.
19. Myrdal, P.; Ward, G. H.; Simamora, P.; Yalkowsky, S. H. AQUAFAC: Aqueous Functional Group Activity Coefficients. *SAR QSAR Environ. Res.* **1993**, *1*, 53–61.
20. Myrdal, P.; Manka, A. M.; Yalkowsky, S. H. AQUAFAC 3: Aqueous Functional Group Activity Coefficients; Application to the Estimation of Aqueous Solubility. *Chemosphere* **1995**, *30*, 1619–1637.
21. Lee, Y. C.; Myrdal, P.; Yalkowsky, S. H. Aqueous Functional Group Activity Coefficients (AQUAFAC 4): Application to Complex Organic Compounds. *Chemosphere* **1996**, *33*, 2129–2144.
22. Pinsuwan, S.; Myrdal, P.; Lee, Y.C.; Yalkowsky, S. H. AQUAFAC 5: Aqueous Functional Group Activity Coefficients; Application to Alcohols and Acids. *Chemosphere* **1997**, *35*, 2503–2513.

23. Dannenfelser, R. M.; Yalkowsky, S. H. Estimation of Entropy of Melting from Molecular Structure: A Non-Group Contribution Methodol. *Ind. Eng. Chem. Res.* **1996**, *35*, 1483–1486.
24. Simamora, P.; Miller, A. H.; Yalkowsky, S. H. Melting Point and Normal Boiling Point Correlations: Applications to Rigid Aromatic Compounds. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 437–440.
25. Myrdal, P. B.; Yalkowsky, S. H. Estimating Pure Component Vapor Pressures of Complex Organic Molecules. *Ind. Eng. Chem. Res.* **1997**, *36*, 2494–2499.
26. Mackay, D.; Bobra, A.; Chan, D. W.; Shiu, W. Y. Vapor Pressure Correlations for Low-Volatility Environmental Compounds. *Environ. Sci. Technol.* **1982**, *16*, 645–649.
27. Yalkowsky, S. H.; Mishra, D. S. Ideal Solubility of a Solid Solute: Effect of Heat Capacity Assumptions *Pharm. Res.* **1992**, *9*, 958–959.
28. Yalkowsky, S. H.; Mishra, D. S. Estimation of Vapor Pressure of Some Organic Compounds. *Ind. Eng. Chem. Res.* **1991**, *30*, 1609–1612.
29. Hildebrand, J. H.; Scott, R. L. *The Solubility of Nonelectrolytes*; Reinhold Publishing Corporation: New York, 1950.
30. Lide, D. R., Ed. *Handbook of Chemistry and Physics*, 72nd ed.; CRC Press: Boca Raton, FL, 1991–1992.
31. Gross, W.; Keilhorn, J.; Melber, C. *Polybrominated Biphenyls*; World Health Organization: Geneva, 1994.
32. *Aldrich Catalog of Fine Chemicals*; Aldrich Chemical Co.: Milwaukee, WI, 1998–1999.
33. Mackay, D.; Shiu, W. Y.; Ma, K. C. *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*; Lewis Publishers: Ann Arbor, MI, 1992.
34. Horvath, A. L. *Halogenated Hydrocarbons Solubility-Miscibility with Water*; Marcel Dekker: New York, 1982.
35. Yalkowsky, S. H.; Dannenfelser, R. M. *Arizona Database*, 5th ed.; College of Pharmacy, University of Arizona: Tuscon, AZ, 1990.
36. Ishihara, Y.; Takano, J.; Yasuoka, T.; Mitsuzawa, S. The Solubilities of Aromatic Nitro Compounds in Pure water. *J. Chem. Soc. Jpn.* **1996**, *11*, 987–990.
37. Baum, E. J. *Chemical Property Estimation Theory and Application*; Lewis Publishers: Boca Raton, FL, 1998.
38. Pinsuwan, S.; Li, A.; Yalkowsky, S. H. Correlation of Octanol/Water Solubility Ratios and Partition Coefficients. *J. Chem. Eng. Data* **1995**, *40*, 623–626.
39. Verschuere, K. *Handbook of Environmental Data on Organic Chemicals*; Van Nostrand Reinhold Co.: New York, 1977.
40. Myrdal, P. B. A Simple Scheme for Estimating Environmentally Relevant Physical Properties of Organic Compounds. Ph.D. Thesis. Department of Pharm. Sci., College of Pharmacy, University of Arizona, Tuscon, AZ, 1994.
41. Sacan, M. T.; Balcioglu, I. A. Estimation of Liquid Vapor Pressures for Low Volatility Chemicals. *Chemosphere* **1998**, *36*, 451–460.
42. Harner, T.; Bidleman, T. F. Measurement of Octanol-Air Partition Coefficients for Polycyclic Aromatic Hydrocarbons and Polychlorinated Naphthalenes. *J. Chem. Eng. Data* **1998**, *43*, 40–46.
43. Heekman, D. ClogP for Windows version 2.0.0b, 1995–1997, Biobyte Corp., Claremont, CA.

JS990117P