

## A Simple QSPR Model for Predicting Soil Sorption Coefficients of Polar and Nonpolar Organic Compounds from Molecular Formula

Eduardo J. Delgado,\* Joel B. Alderete, and Gonzalo A. Jaña

Theoretical and Computational Chemistry Group (QTC), Faculty of Chemical Sciences,  
Universidad de Concepción, Concepción, Chile

Received August 1, 2003

A quantitative structure–property relationship (QSPR) model is developed to predict the logarithm of the soil sorption coefficient of 82 organic compounds. The data set contains polar and nonpolar, saturated, unsaturated, aliphatic, aromatic, and polycyclic aromatic compounds covering a log  $K_{oc}$  range from about 1 to 6 log units. The best correlation equation, containing only five constitutional descriptors (number of benzene rings, molecular weight, number of N, O, and S atoms), predicts log  $K_{oc}$  with a squared correlation coefficient of 0.94, having a standard deviation,  $s$ , of 0.33. The model is validated with an external set of 43 compounds not included in the training set. The descriptors involved in the model can be obtained easily from the molecular formula without any further calculation; therefore, the model is ready to use by environmental scientists with no background in quantum chemistry or chemical graph theory or when no software is available.

### INTRODUCTION

The term sorption is used frequently in environmental situations to denote the uptake of a solute by a solid (soil or sediment or component of soil) without reference to a specific mechanism. Sorption processes play a major role in determining the environmental fate and impact of organic chemicals. Sorption affects a variety of specific fate processes, including volatilization,<sup>1</sup> bioavailability,<sup>2</sup> biodegradability,<sup>3</sup> photolysis,<sup>4</sup> and hydrolysis.<sup>5</sup>

Sorption coefficients,  $K_{oc}$ , quantitatively describe the extent to which an organic chemical is distributed at equilibrium between an environmental solid, soil, sediment, suspended sediment, and the aqueous phase it is in contact with at equilibrium

$$K_{oc} = \frac{C_{Soil}}{C_w} \quad (1)$$

where  $C_{Soil}$  is the concentration of solute per gram of carbon in the soil phase, and  $C_w$  denotes the concentration of solute in the aqueous phase.  $K_{oc}$  plays an important role in describing the pollution potential of a nonionic organic pesticide, such pollution being mainly due to the pesticides undergoing leaching or runoff when applied to the soil.

Sorption coefficients generally are determined from an isotherm, a diagram that depicts the distribution of the test chemical between a solid sorbent and the solution in equilibrium with it over a range of concentrations at constant temperature. However, the experimental measurement of  $K_{oc}$  is expensive, time-consuming, and often related with considerable experimental error; consequently, there is a great need for reliable calculation methods which can be used for the prediction of  $K_{oc}$ . Many methods have been reported which are based on correlations with other experimental data,

especially with octanol/water partition coefficients, water solubilities, and melting points.<sup>6</sup> Several other methods have been developed which do not rely on experimental data. Some of these models require only topological descriptors, like molecular connectivity indices<sup>7,8</sup> or the so-called adsorbability index. Another approach is based on fragment contributions; for example Meylan et al.<sup>9</sup> developed a model combining topological and fragment contributions. The polar correction factors assume that a single numerical value represents the contribution of each polar functional group to log  $K_{oc}$  and that the contributions made by each group are independent of each other. This method appears to have considerable predictive power. However, the above assumptions make its use somewhat troublesome due to the values of the fragment contributions which depend on the following: the kind of atoms of the molecule, the fragment to which they are bonded to, etc. Moreover, it cannot be applied for compounds with polar fragments, for which no group contributions have been fitted before. Thus it is not applicable to compounds with new polar groups.

Other models have been developed to predict  $K_{oc}$  from molecular properties that are obtained from quantum chemical calculations. Recently, Winget et al.<sup>10</sup> reported a study in which they predict  $K_{oc}$  using a universal solvation model, SMx, which is based on semiempirical molecular orbital theory in combination with a dielectric continuum model. Klamt et al.<sup>11</sup> reported the prediction of  $K_{oc}$  using another universal solvation model, the conductor-like screening model for real solvent (COSMO-RS) which is based on density functional theory.

Since all the above methods, based either on other experimental properties or topological parameters in conjunction with fragment contributions, or quantum chemical calculations, require information and/or computer programs which are not always readily available, it would be greatly desirable to have a model based on parameters obtained

\* Corresponding author e-mail: edelgado@udec.cl.

directly from the molecular formula without any further calculation. A model with this characteristic would be of great importance for environmental scientists with no background in quantum chemistry or chemical graph theory.

In this article we report a QSPR model for the prediction of  $K_{oc}$  based only on five constitutional descriptors which can be obtained easily from the molecular formula simply by counting.

### CHEMICAL DATA

The data set of the sorption coefficients was collected from several literature sources.<sup>12–14</sup> A total of 82 structurally diverse compounds were considered. The data set contains both polar and nonpolar, saturated, unsaturated, aliphatic, aromatic, and polycyclic aromatic compounds covering a log  $K_{oc}$  range from about 1 to 6 log units.

### COMPUTATIONAL METHODS

The quantitative structure–property relationship was developed using the Microsoft Windows version of the Codessa program,<sup>15</sup> which is a chemical multipurpose quantitative-activity and structure–property statistical analysis and prediction program, using the best correlation option. This program calculates a total of 38 constitutional descriptors, including the following: absolute and relative numbers of atoms, number of single, double, triple, and aromatic bonds, number of rings, number of benzene rings, and molecular weight.

### RESULTS AND DISCUSSION

For the 82 compounds the best correlation equation involving five constitutional descriptors is

$$\text{Log } K_{oc} = 0.51 + 0.60 N_{\phi} + 1.02 \times 10^{-2} M_W - 0.48 N_N - 0.25 N_O + 0.61 N_S$$

where  $N_{\phi}$  is the number of benzene rings,  $M_W$  is the molecular weight,  $N_N$  is the number of nitrogen atoms,  $N_O$  is the number of oxygen atoms, and  $N_S$  is the number of sulfur atoms. Summary statistics for the correlation of experimental vs calculated log  $K_{oc}$  for the 82 compounds are

$$R^2 = 0.94; (R_{CV})^2 = 0.93; s = 0.33; F = 227.51$$

where  $R$  is the correlation coefficient,  $F$  is the value of the Fisher test,  $s$  is the standard deviation of the fit, and  $R_{CV}$  is the cross-validated correlation coefficient. This coefficient describes the stability of the regression model obtained by focusing on the sensitivity of the model to the elimination of any single data point. Briefly, for each data point, the regression is recalculated with the same descriptors but for the data set without this point. The obtained regression is used to predict the value of this point, and the set of estimated values calculated in this way is correlated with the experimental values.

Table 1 summarizes the results of the best five parameter QSPR model derived for log  $K_{oc}$ . The first column contains the constitutional descriptors involved in the correlation; the second column contains the respective values of the regres-

**Table 1.** Best Five Descriptors Correlation Model of log  $K_{oc}$

descriptor	coefficient	<i>t</i> -test
intercept	$0.51 \pm 0.16$	3.26
number of benzene rings	$0.60 \pm 3.65 \times 10^{-2}$	16.54
molecular weight	$1.01 \times 10^{-2} \pm 5.65 \times 10^{-4}$	17.99
number of N atoms	$-0.48 \pm 6.84 \times 10^{-2}$	-6.95
number of O atoms	$-0.25 \pm 5.12 \times 10^{-2}$	-4.95
number of S atoms	$0.61 \pm 0.11$	5.36

**Table 2.** Correlation Matrix of the Best Five Descriptors

	A <sup>a</sup>	B <sup>b</sup>	C <sup>c</sup>	D <sup>d</sup>	E <sup>e</sup>
A <sup>a</sup>	1.000				
B <sup>b</sup>	0.078	1.000			
C <sup>c</sup>	-0.249	-0.168	1.000		
D <sup>d</sup>	-0.183	0.009	0.739	1.000	
E <sup>e</sup>	-0.387	-0.077	0.218	0.162	1.000

<sup>a</sup> A – number of benzene rings. <sup>b</sup> B – molecular weight. <sup>c</sup> C – number of N atoms. <sup>d</sup> D – number of O atoms. <sup>e</sup> E – number of S atoms.

sion coefficients along with their mean square errors; and finally the third column contains the *t*-student values for the correlation coefficients; these values reflect the significance of each parameter within the model. This test indicates that the most important descriptors are the molecular weight and the number of benzene rings.

According to statistics, a regression equation is of no relevance when the variables are mutually interrelated by simple or multiple correlation, i.e., they are nonorthogonal. This possibility is excluded in this study as it can be inferred from the elements of the correlation matrix, Table 2.

Positive values in the regression coefficients indicate that those descriptors contribute positively to the value of the soil sorption coefficient, whereas negative values indicate that the greater the value of these descriptors the lower the value of  $K_{oc}$ .

The calculated and experimental values of the logarithm of the soil sorption coefficient are compared in Table 3, and the scatter plot is shown in Figure 1.

To check the predictive capability of the model, it was tested with an external set of chemicals not included in the training set. The validation data set included 43 compounds with a diverse selection of chemical structures. It contained not only many chemicals that are similar in structure to chemicals in the training set but also some chemicals having other functional groups, namely, alcohols and carboxylic acids.

Table 4 lists the experimental and calculated log  $K_{oc}$  values for the validation set. Statistical performance for calculated vs experimental log  $K_{oc}$  for the validation was as follows:  $R = 0.96$ ,  $R^2 = 0.91$ ,  $s = 0.09$ ,  $s^2 = 0.30$ . These results confirm the predictive capability of the model.

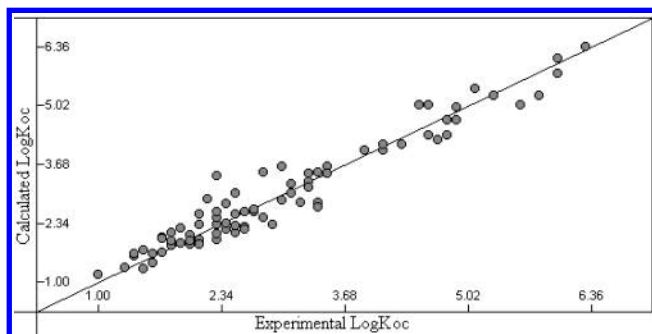
### PHYSICAL INTERPRETATION

According to the definition of  $K_{oc}$ , eq 1, it is evident that its value is determined by the concentration of solute in the soil phase, i.e., by the sorption process, and the concentration in the aqueous phase, i.e., by the solubility process.

Sorption occurs when the free energy of the interaction between an environmental solid and an organic chemical sorbate is negative. This process is driven by relatively weak intermolecular forces called van der Waals type interactions.

**Table 3.** Experimental and Calculated log  $K_{oc}$  Values

compound	CAS no.	exp	calc	diff	compound	CAS no.	exp	calc	diff
benzene	71-43-2	1.8	1.9	-0.1	propham	122-42-9	2.3	2.0	0.3
naphthalene	91-20-3	3.1	3.0	0.1	asulam	3337-71-1	2.5	2.1	0.4
anthracene	120-12-7	4.3	4.1	0.2	carbaryl	63-25-2	2.4	2.8	-0.4
tetracene	92-24-0	5.8	5.2	0.6	carbofuran	1563-66-2	1.8	2.1	-0.3
pyrene	129-00-0	4.9	5.0	-0.1	diallate(cis)	2303-16-4	3.3	3.1	0.2
benzo[a]pyrene	50-32-8	6.0	6.1	-0.1	methomyl	16752-77-5	1.3	1.3	0.0
benz[a]anthracene	56-55-3	5.3	5.2	0.1	oxamyl	23135-22-0	1.0	1.2	-0.2
dibenz[a,h]anthracene	53-70-3	6.3	6.4	-0.1	pebulate	1114-71-2	2.8	2.5	0.3
1,1,1-trichloroethane	71-55-6	2.0	1.9	0.1	methyl-n-phenylcarbamate	2603-10-3	1.7	1.7	0.0
1,1,2-trichloroethane	79-00-5	1.9	1.9	0.0	ethyl-n-phenylcarbamate	101-99-5	1.8	1.8	0.0
1,1,2,2-tetrachloroethane	79-34-5	1.9	2.2	-0.3	propyl-n-phenylcarbamate	5532-90-1	2.1	2.0	0.1
trichloroethene	79-01-6	2.0	1.8	0.2	butyl-n-phenylcarbamate	1538-74-5	2.3	2.1	0.2
tetrachloroethene	127-18-4	2.4	2.2	0.2	pentyl-n-phenylcarbamate	63075-06-9	2.6	2.2	0.4
biphenyl	92-52-4	3.3	3.3	0.0	propoxun	114-26-1	1.7	2.0	-0.3
2-chlorobiphenyl	2051-60-7	3.5	3.6	-0.1	triallate	2303-17-5	3.4	3.5	-0.1
2,2'-dichlorobiphenyl	13029-08-8	3.9	4.0	-0.1	vernolate	1929-77-7	2.3	2.5	-0.2
2,4'-dichlorobiphenyl	34883-43-7	4.1	4.0	0.1	chloroxuron	1982-47-4	3.1	3.2	-0.1
2,2',4-PCB	37680-66-3	4.8	4.3	0.5	diuron	33-54-1	2.5	2.3	0.2
2,2',5-PCB	37680-65-2	4.6	4.3	0.3	linuron	330-55-2	2.6	2.2	0.4
2,4,4'-PCB	7012-37-5	4.6	4.3	0.3	monolinuron	1746-81-2	2.1	1.8	0.3
2,5,2',5'-PCB	35693-99-3	4.9	4.7	0.2	neburon	555-37-3	3.4	2.7	0.7
2,4,5,2',5'-PCB	37680-73-2	4.6	5.0	-0.4	metoxuron	19937-59-8	1.7	2.0	-0.3
2,3,4,2',5'-PCB	38380-02-8	4.5	5.0	-0.5	chlorotoluron	15545-48-9	2.0	2.1	-0.1
2,4,5,2',4',5'-PCB	35065-27-1	5.6	5.0	0.6	metobromuron	3060-89-7	2.1	2.3	-0.2
2,3,4,2',3',4'-PCB	38380-07-3	5.1	5.4	-0.3	phenylurea	64-10-8	1.5	1.3	0.2
2,3,4,5,6,2',5'-PCB	52712-05-7	6.0	5.7	0.3	2-chlorophenylurea	114-38-5	1.6	1.6	0.0
$\alpha$ -benzenehexachloride	319-84-6	3.3	3.5	-0.2	3-methylphenylurea	63-99-0	1.6	1.4	0.2
$\beta$ -benzenehexachloride	319-85-7	3.5	3.5	0.0	fenuron	101-42-8	1.4	1.6	-0.2
$\gamma$ -benzenehexachloride	58-89-9	3.3	3.5	-0.2	<i>N</i> -(3-chlorophenyl)- <i>N'</i> , <i>N'</i> -dimethylurea	587-34-8	1.8	1.9	-0.1
aldrin	309-00-2	4.7	4.2	0.5	monuron	150-68-5	2.0	1.9	0.1
endrin	72-20-8	4.1	4.1	0.0	<i>N</i> -(4-methylphenyl)- <i>N'</i> , <i>N'</i> -dimethylurea	7160-01-2	1.5	1.7	-0.2
dieldrin	60-57-1	4.1	4.1	0.0	<i>N</i> -(4-methoxyphenyl)- <i>N'</i> , <i>N'</i> -dimethylurea	7160-02-3	1.4	1.6	-0.2
chlordane	57-74-9	4.8	4.7	0.1	chlorbromuron	13360-45-7	2.7	2.6	0.1
butylate	2008-41-5	2.7	2.6	0.1	propyzamide	23950-58-5	3.0	2.9	0.1
chlorpropham	101-21-3	2.9	2.3	0.6	alachlor	15972-60-8	2.2	2.9	-0.7
cycloate	1134-23-2	2.6	2.6	0.0	isoxaben	82558-50-7	2.5	2.6	-0.1
desmediphan	13684-56-5	3.2	2.8	0.4	metolachlor	51218-45-2	2.5	3.0	-0.5
eptc	759-94-4	2.4	2.3	0.1	propachlor	1918-16-7	2.1	2.5	-0.4
molinate	2212-67-1	2.3	2.3	0.0	propanil	709-98-8	2.3	2.6	-0.3
phenmedhipam	13684-63-4	3.4	2.8	0.6	diphenamid	957-51-7	2.3	3.4	-1.1
thiobencarb	28249-77-6	3.0	3.6	-0.6	napropamide	15299-99-7	2.8	3.5	-0.7

**Figure 1.** Scatter plot of the calculated vs experimental log  $K_{oc}$ .

This type of interaction refers to all the attractive and repulsive interaction between chemically nonbonding molecules, namely, electrostatic, induction, and dispersion interaction. However, the dominant attractive interaction between uncharged species is the dispersion force, which is also called the London force. The dispersion forces arise from the coupling of instantaneous fluctuations in the electron density on two neighboring molecules.

Instantaneous dipoles generated in a molecule, as a result of these fluctuations, induces dipoles in neighboring molecules, and provided the orientations of the two are appropriate, there will be an attractive interaction between them. It is expected that molecules containing  $\pi$  electrons, due to

the special mobility of these electrons, will result in an enhanced polarizability and interaction energy of unsaturated molecules with the substrate favoring in this way the sorption process and therefore leading to an increase in the value of  $K_{oc}$ .

On the other hand, the solubility is almost exclusively dependent on the intermolecular forces that exist between the solute molecules and the solvent molecules. Therefore the solute-solvent interactions, namely, hydrogen bonding, dipole-dipole interactions, are largely responsible for the amount of the compound dissolved. Those forces are dependent on the composition and structure of the compound. The introduction of substituents into organic compounds with increasing differences in electronegativity with respect to carbon, e.g. nitrogen or oxygen atoms, should favor the solute-solvent interaction and therefore enhancing the solubility of the compound and consequently decreasing the values of  $K_{oc}$ .

In the present model, the number of benzene rings and molecular weight are the most important descriptors, as can be concluded from its *t*-test value. The positive values of the respective correlation coefficients for these descriptors indicate the value of  $K_{oc}$  goes up as the number of benzene rings and molecular weight increase. This behavior may be explained as follows.

**Table 4.** Experimental and Calculated log  $K_{oc}$  Values for the Validation Set

compound	CAS no.	exp	calc	diff
3-chloro-4-methoxyphenylurea	25277-05-8	2	1.7	0.3
3-methyl-4-bromophenylurea	78508-46-0	2.4	2.2	0.2
4-bromophenylurea	1967-25-5	2.1	2.1	0.0
methyl- <i>N</i> -(3-chlorophenyl)carbamate	2150-88-1	2.2	2.0	0.2
methyl- <i>N</i> -(3,4-dichlorophenyl)carbamate	1918-18-9	2.7	2.4	0.3
<i>N</i> -(3,5-dimethylphenyl)- <i>N'</i> , <i>N'</i> -dimethylurea	36627-56-2	1.7	1.9	-0.2
<i>N</i> - <i>N</i> -(3,5-dimethyl-4-bromophenyl)- <i>N'</i> , <i>N'</i> -dimethylurea	78508-43-7	2.5	2.7	-0.2
<i>N</i> -(3-fluorophenyl)- <i>N'</i> , <i>N'</i> -dimethylurea	330-39-2	1.7	1.8	-0.1
<i>N</i> -(3-chloro-4-methylphenyl)- <i>N'</i> -methylurea	22175-22-0	2.1	1.9	0.2
<i>N</i> -(3-chloro-4-methoxyphenyl)- <i>N'</i> -methylurea	20782-57-4	1.8	1.8	0.0
<i>N</i> -phenyl- <i>N</i> -methylurea	4559-87-9	1.3	1.4	-0.1
<i>N</i> -phenyl- <i>N'</i> -cyclohexylurea	86759-64-0	2.1	2.1	0.0
<i>N</i> -phenyl- <i>N'</i> -cyclopropylurea	13140-86-8	1.7	1.7	0.0
<i>N</i> -phenyl- <i>N'</i> -cyclopentylurea	13140-89-1	1.9	2.0	-0.1
bensulfuron-methyl	83055-99-6	2.6	2.2	0.4
thifensulfuron-methyl	79277-27-3	1.6	1.8	-0.2
metsulfuron-methyl	74223-64-6	1.5	1.7	-0.2
4-phenoxyphenylurea	78508-44-8	2.6	2.6	0.0
3-methyl-4-fluorophenylurea	78508-45-9	1.7	1.5	0.2
1,2,3,4-tetrachlorobenzene	634-66-2	3.8	3.3	0.5
<i>o</i> -xylene	95-47-6	2.1	2.0	0.1
<i>N,N</i> -dimethylbenzamide	611-74-5	1.6	1.9	-0.3
<i>N</i> -methylbenzamide	613-93-4	1.8	1.8	0.0
1,2,3,5-tetrachlorobenzene	634-90-2	3.2	3.3	-0.1
trichlorobenzene	12002-48-1	3.6	3.0	0.6
trimethylbenzene	25551-13-7	2.8	2.3	0.5
dichlorobenzene	25321-22-6	2.5	2.6	-0.1
dimethylbenzene	95-47-6	2.5	2.2	0.3
tetrachlorophenol	25167-83-3	3.5	3.2	0.3
dichlorophenol	25167-81-1	2.5	2.5	0.0
trichlorophenol	25167-82-2	3.1	2.9	0.2
methylnaphthalene	1321-94-4	3.9	3.2	0.7
chlorobenzene	108-90-7	2.4	2.3	0.1
hexachlorobenzene	118-74-1	3.6	4.0	-0.4
pentachlorobenzene	108-93-5	4.1	3.7	0.4
toluene	108-88-3	2.1	2.1	0.0
trichloromethane	67-66-3	1.6	1.7	-0.1
tetrachloromethane	56-23-5	1.8	2.1	-0.3
hexanoic acid	142-62-1	1.5	1.2	0.3
benzoic acid	65-85-0	1.5	1.9	-0.4
methanol	67-56-1	0.4	0.6	0.2
1-hexanol	111-27-3	1.0	1.3	-0.3
1-heptanol	111-70-6	1.1	1.4	-0.3

The number of benzene rings encode, on one hand, the hydrophobicity of the compound, i.e., its tendency to exclude from water. Thus, an increase in this descriptor leads to an decrease in the solubility of the compound. On the other hand, the special mobility of  $\pi$  electrons will result in an enhanced polarizability and interaction energy of unsaturated molecules with the solid and therefore favoring the sorption process. Both phenomena, surely operating simultaneously, lead to an increase in the value of  $K_{oc}$ .

The molecular weight affects mainly the solubility of the compound as it is well-known in physical organic chemistry, namely, the larger the molecular weight the lower the solubility of the compound, leading to an increase in the value of  $K_{oc}$  due to its inversely dependence on solubility.

The correlation coefficients for the number of nitrogen and oxygen atoms are negative indicating that the greater the value of these descriptors the lower the value of  $K_{oc}$ . This behavior can be explained because the presence of these atoms enhances the solubility of the compound due to the hydrogen bonding and dipole-dipole interactions as it was explained above.

The correlation coefficient for the number of sulfur atoms is positive indicating an increase in the value of this

descriptor leads to an increase in the value of  $K_{oc}$ . The negligible difference of electronegativity between sulfur and carbon, about 0.03 units in the Pauling scale, leads to relatively small bond dipoles disfavoring the solute-solvent interactions and therefore decreasing the solubility and consequently increasing the value of  $K_{oc}$ .

## CONCLUSIONS

The merit of the QSPR model developed in this article lays in its simplicity. Soil sorption coefficients can be predicted straightforwardly from the molecular formula, for both nonpolar and polar compounds, without any calculation of indices or more complicated quantum chemical calculations or without the need of having the values of polar fragment contributions. Therefore the model is applicable to compounds with polar fragments for which no group contributions have been fitted before. Currently, the model is being developed to include organophosphorus aliphatic and aromatic compounds.

## ACKNOWLEDGMENT

The authors acknowledge financial support from Fondecyt under Grant No. 1020464.



## REFERENCES AND NOTES

- (1) Chiou, C. T.; Shoup, T. D. Soil sorption of organic vapors and effects of humidity on sorption mechanism and capacity. *Environ. Sci. Technol.* **1985**, *19*, 1196–1200.
- (2) Ditoro, D. M.; Zarba, C. S.; Hasen, D. J.; Berry, W. J.; Swartz, R. C.; Cowan, C. E.; Pavlov, S. P.; Allen, H. E.; Thomas, N. A.; Paquin, P. R. Technical basis for establishing sediment quality criteria for nonionic organic chemicals using equilibrium partitioning. *Environ. Toxicol. Chem.* **1991**, *10*, 1541–1583.
- (3) Scow, K. M.; Johnson, C. Effect of sorption on biodegradation of soil pollutants. In *Advance in Agronomy*; Sparks, D. L., Ed.; Academic: San Diego, 1997; pp 1–56.
- (4) Tao, S.; Piao, H. S.; Dawson, R.; Lu, X. X.; Hu, H. Y. Estimation of organic carbon normalized sorption coefficient ( $K_{oc}$ ) for soils using the fragment constant method. *Environ. Sci. Technol.* **1999**, *33*, 2719–2725.
- (5) Deeley, G. M.; Reinhard, M.; Stearns, S. M. Transformation and sorption of 1,2-dibromo-3-chloropropane in subsurface samples collected at Fresno, California. *J. Environ. Qual.* **1991**, *20*, 547–556.
- (6) Boethling, R. S.; Mackay, D. *Handbook of property estimation methods for chemicals: environmental and health sciences*; Lewis Publishers: Boca Raton, 2000; pp 141–188.
- (7) Sabljic, A. Predictions of the nature and strength of soil sorption of organic pollutants by molecular topology. *J. Agric. Food Chem.* **1984**, *32*, 243–246.
- (8) Sabljic, A. On the prediction of soil sorption coefficients of organic pollutants from molecular structure: application of molecular topology model. *Environ. Sci. Technol.* **1987**, *21*, 358–366.
- (9) Meylan, W.; Howard, P. H.; Boethling, R. S. Molecular topology/fragment contribution method for predicting soil sorption coefficients. *Environ. Sci. Technol.* **1992**, *26*, 1560–1567.
- (10) Winget, P.; Cramer, C. J.; Truhlar, D. G. Prediction of soil sorption coefficients using a universal solvation model. *Environ. Sci. Technol.* **2000**, *34*, 4733–4740.
- (11) Klamt, A.; Eckert, F.; Diedenhofen, M. Prediction of soil sorption coefficients with a conductor-like screening model for real solvent. *Environ. Toxicol. Chem.* **2002**, *21*, 2562–2566.
- (12) Syracuse Research Corporation, *Physical/Chemical Property Database (Physptop)*; SRC Environmental Science Center: Syracuse, NY, 1994.
- (13) Linders, J. B. H. J.; Jansma, J. W.; Mensik, B. J. W. G.; Otermann, K. National Institute of Public Health and Environmental Protection (RIVM), Bilthoven, The Netherlands, Report No. 679101014.
- (14) Mackay, D.; Shiu, W. Y.; Ma, K. C. *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*; Lewis Publishers: Boca Raton, 1997.
- (15) Katritzky, A. R.; Lovanov, V. S.; Karelson, M. *Codessa version 2.0 Reference Manual*, 1994.

CI0341666