

Aqueous Biphasic Systems. Partitioning of Organic Molecules: A QSPR Treatment

Alan R. Katritzky,* Kaido Tamm, Minati Kuanar, and Dan C. Fara

Center for Heterocyclic Compounds, Department of Chemistry, University of Florida,
Gainesville, Florida 32611

Alexander Oliferenko and Polina Oliferenko

Department of Chemistry, Moscow State University, Moscow 119899, Russia

Jonathan G. Huddleston and Robin D. Rogers

Department of Chemistry and Center for Green Manufacturing, The University of Alabama,
Tuscaloosa, Alabama 35487

Received September 2, 2003

The partitioning of 29 small organic probes in a PEG-2000/(NH₄)₂SO₄ biphasic system was investigated using a quantitative structure–property relationship (QSPR) approach. A three-descriptor equation with the squared correlation coefficient (R^2) of 0.97 for the partition coefficient ($\log D$) was obtained. All descriptors were derived solely from the chemical structure of the compounds. Using the same descriptors, a three-parameter model was also obtained for $\log P$ (octanol/water, $R^2 = 0.89$); predicted $\log P$ values were used as an external descriptor for modeling $\log D$.

INTRODUCTION

Aqueous biphasic systems (ABS) are formed by the addition of two (or more) water-soluble polymers or a polymer and salt to an aqueous solution above certain critical concentrations or temperature. ABS are unique because each of the two nonmiscible phases is over 80% water on a molal basis and possesses different solvent properties.^{1,2} Due to their highly aqueous and hence mild nature, which is consonant with the maintenance of macromolecular structure, ABS have been employed for the separation of biological macromolecules for over 40 years.^{2,3} ABS systems are also used in industrial biotechnology quality control for the detection of denaturation and degradation of proteins.⁴ ABS partition coefficients have been suggested as viable alternatives to $\log P$ values in QSAR (quantitative structure–activity relationships) applied to biotechnological products.^{4,5} Proteins as biomacromolecules and nucleic acids are prone to denaturation in alkyl alcohols, whereas suitable polymer/water compositions can more closely resemble the native living cell conditions. Researchers studying the partitioning of macromolecular biological solutes also examine the partitioning of small organic molecules to understand and predict macromolecular partitioning in ABS.^{6–8} Eiteman and Gainer have studied the partitioning of amino acids, small peptides and also some alcohols in ABS.^{9,10}

Recently, partitioning in ABS has found applications in many different fields of science and technology due to ABS media being nonvolatile, nontoxic, and nonflammable. ABS thus represent unique alternatives to traditional solvent based biphasic systems for the separation of metal ion species,^{3,11–14} small organic molecules,^{15,16} and lignins from cellulose in the paper and pulping process.^{17,18}

Experimental investigations of the partitioning behavior of organic molecules in poly(ethylene glycol)-salt ABS have been reported by Rogers et al.^{15,16} The partition coefficient (D) is defined (eq 1) as the total concentration of a solute in the upper PEG-rich phase divided by the concentration in the lower salt-rich phase. In eq 1, C_{PEG} and C_{Salt} are the solute concentrations in the corresponding phases.^{15,16} Usually the logarithmic function ($\log D$) is used for describing the partition coefficients.

$$D = \frac{C_{\text{PEG}}}{C_{\text{Salt}}} \quad (1)$$

The partition of several relatively hydrophilic substituted charged and uncharged benzene species can be defined in terms of the free energy of hydration of the salt forming the ABS and the dissociation of the charged groups of the distributed solutes. For a given solute and a defined concentration of salt, the partition coefficients were observed to increase as the hydration free energy value of the salt anion becomes more negative, and, in a given salt, the $\log D$ increases in the order benzene < toluene < chlorobenzene < 1,4-dichlorobenzene < 1,2,4-trichlorobenzene. The partitioning behavior of the solutes in PEG-salt ABS was also found to be strongly correlated with their partitioning coefficients in 1-octanol–water biphasic systems.¹⁵ In continuation of the study of partitioning in ABS, Rogers et al. studied the relative hydrophobicities of several PEG/salt ABSs composed of different molecular weights of PEG (1000, 2000, and 3400) and a variety of inorganic salts (K₃-PO₄, K₂CO₃, (NH₄)₂SO₄, Li₂SO₄, MnSO₄, ZnSO₄, and NaOH).^{16,19} The $\log D$ was shown to be a function only of the degree of phase divergence of the biphasic system as expressed by the difference in polymer concentration between

* Corresponding author phone: (352)392-0554; fax: (352)392-9199; e-mail: katritzky@chem.ufl.edu.

Table 1. Observed and Predicted log *D* Values and Descriptors Used in Study

no.	name	experimental ^a		predicted			descriptors used in study ^b						
		log <i>D</i>	log <i>P</i>	eq 2 log <i>D</i>	eq 7 log <i>P</i>	eq 8 log <i>D</i>	<i>G</i> _b	SE _{OK}	<i>q</i> _{min}	⁰ <i>χ</i> ^v	¹ <i>χ</i> ^v	² <i>χ</i> ^v	³ <i>χ</i> ^v
1	benzene	1.48	2.13	1.37	2.23	1.54	504.76	0.00	-0.13	3.46	2.00	1.16	0.67
2	toluene	1.55	2.73	1.46	2.31	1.58	588.94	0.00	-0.18	4.39	2.41	1.66	0.94
3	1,2-dichloroethane	0.90	1.48	1.08	1.73	1.28	379.04	0.00	-0.13	3.68	2.10	1.13	0.64
4	ethyl acetate	0.57	0.73	0.88	0.93	0.86	527.98	0.04	-0.36	4.02	1.90	0.93	0.35
5	methanol	0.10	-0.77	-0.03	-0.70	0.02	142.96	0.08	-0.33	1.45	0.45	0.00	0.00
6	ethanol	0.25	-0.31	0.21	-0.21	0.27	224.20	0.05	-0.33	2.15	1.02	0.32	0.00
7	<i>n</i> -propanol	0.49	0.25	0.42	0.19	0.48	306.49	0.04	-0.33	2.86	1.52	0.72	0.22
8	2-propanol	0.40	0.05	0.41	0.16	0.47	305.17	0.04	-0.33	3.03	1.41	1.09	0.00
9	<i>n</i> -butanol	0.68	0.88	0.63	0.57	0.68	388.67	0.03	-0.33	3.57	2.02	1.08	0.51
10	<i>n</i> -pentanol	0.92	1.51	0.83	0.92	0.86	470.88	0.03	-0.33	4.28	2.52	1.43	0.76
11	<i>n</i> -octanol	1.54	3.00	1.41	1.95	1.39	717.49	0.02	-0.33	6.40	4.02	2.49	1.51
12	chlorobenzene	1.81	2.84	1.65	2.68	1.77	641.38	0.02	-0.13	4.52	2.48	1.73	0.99
13	acetonitrile	0.23	-0.34	0.28	-0.23	0.26	222.88	0.20	-0.15	1.95	0.72	0.22	0.00
14	aniline	1.08	0.90	1.14	1.47	1.14	606.29	0.03	-0.33	3.96	2.20	1.41	0.80
15	acetophenone	1.38	1.58	1.50	2.00	1.42	778.46	0.07	-0.30	5.30	2.87	1.92	1.18
16	4-hydroxybenzoic acid	1.63	1.58	1.68	2.22	1.53	914.22	0.07	-0.37	5.11	2.72	1.85	1.11
17	benzyl alcohol	1.16	1.10	1.35	1.84	1.34	690.63	0.02	-0.33	4.54	2.58	1.64	1.03
18	acetic acid	0.23	-0.17	0.40	0.04	0.40	340.92	0.06	-0.36	2.36	0.93	0.52	0.00
19	benzoic acid	1.45	1.87	1.48	1.92	1.38	806.36	0.05	-0.37	4.74	2.59	1.67	1.02
20	nitrobenzene	1.36	1.85	1.30	1.15	0.98	879.27	0.22	-0.36	4.65	2.50	1.59	0.97
21	methyl iodide	0.75	1.51	0.56	0.43	0.60	392.21	0.01	-0.38	3.54	2.54	0.00	0.00
22	anisole	1.53	2.11	1.66	2.56	1.71	718.18	0.01	-0.21	4.80	2.52	1.52	0.98
23	phenol	1.32	1.46	1.33	1.93	1.38	612.16	0.02	-0.25	3.83	2.13	1.34	0.76
24	1,4-dichlorobenzene	2.12	3.44	2.02	3.37	2.13	778.57	0.00	-0.12	5.58	2.96	2.31	1.31
25	1,2,4-trichlorobenzene	2.14	4.02	2.33	3.88	2.39	916.84	0.01	-0.12	6.63	3.44	2.81	1.86
26	4,4-dichlorobiphenyl	3.14	5.23	3.02	4.92	2.93	1298.41	0.03	-0.18	9.09	5.26	4.87	2.77
27	<i>p</i> -toluic acid	1.69	2.27	1.65	2.19	1.52	890.46	0.06	-0.37	5.67	3.00	2.17	1.30
28	benzamide	1.24	0.64	1.19	1.14	0.97	800.27	0.11	-0.44	4.87	2.65	1.73	1.06
29	4-chloroaniline	1.44	1.83	1.39	1.81	1.32	743.26	0.05	-0.33	5.02	2.68	1.99	1.12

^a Experimental values from Rogers et al.²⁰ ^b Descriptor's symbols are explained in text.

the phases: Δ[poly-ethylene glycol], Δ[ethylene oxide monomers] or the tie line length (ΔPEG, ΔEO, or TLL, respectively).¹⁹ A linear solvation energy relationship (LSER) based on Abraham's generalized solvation equation was applied, which enabled a direct comparison between the solvent properties of PEG/salt ABS and those of traditional solvent/water systems.^{19,20} Rogers and co-workers have studied the partitioning behavior found in ABS in the context of different applications; their continuing studies in this area explored the partitioning of food coloring dyes²¹ and porphyrin dyes²² in aqueous biphasic systems. The polarities of the PEG-2000-K₃PO₄ aqueous biphasic system have been examined using Reichardt's pyridinium-*N*-phenoxybetaine carboxylate dye as a probe.²³ The solvatochromic properties such as polarity (π^*), H-bond acidity (α), and H-bond basicity (β) have been determined for ABS composed of poly(ethylene glycol) and the salts (NH₄)₂SO₄ or K₃PO₄ using as probes Reichardt's betaine carboxylate, 4-nitrophenol, and 4-nitroanisole. Rogers et al.²⁴ observed that the phase rich in polymer differed little in relative polarity or relative basicity to that rich in salt. However, the polymer rich phase was found to be significantly depleted in H-bond donor ability relative to the salt rich phase and pure water. From a comparative study using a linear free energy relationship (LFER) the Rogers group concluded that the principal molecular determinants, which govern the partitioning in ABS, arise from the size, basicity, and aromaticity or halogenicity of the solute. In a study published in 1999,²⁵ the same group considered the efficacy of ABS as compared to other classical solvent extraction technologies in envi-

ronmentally benign liquid/liquid extraction media, which uses aqueous polymeric solutions. Principal component analysis and nonlinear mapping show that the unconventional partition systems are considerably different from conventional water-organic solvent systems. The use of general LFER analysis of water solvent partitioning systems enables room-temperature ionic liquid (RTIL) and ABS to be compared with conventional systems in general chemical terms.²⁶

Quantitative structure activity/property relationship (QSAR/QSPR) studies are recognized as fast and powerful tools to relate physical and chemical manifestations of compounds to the molecular structures. QSPR treatment of the octanol-water system have been studied by several workers, and various methods of correlating partition coefficients have been reviewed.²⁷ Previously, QSPR methodology based on theoretical molecular descriptors, was applied by the Katritzky group for the prediction of water-air partition coefficients.²⁸ However, it seems that QSPR analysis of ABS is more complex as compared to that of the octanol-water system which has a constant phase composition. ABS are of variable composition, so the QSPR analysis should be carried out not only on one system of some fixed composition but also on several compositions in order to find one possessing desirable properties. No previous theoretical molecular model approach for the prediction of partition in ABS using solely descriptors calculated from chemical structure was located in the literature. The aims and objectives of the present work are to obtain theoretical QSPR models for the description and prediction of solute partitioning in different ABS.

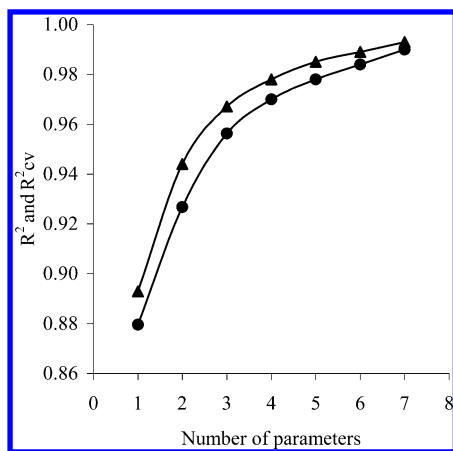


Figure 1. Number of parameters plotted against R^2 (\blacktriangle) and R^2_{cv} (\bullet) values.

DATA SET AND METHODOLOGY

In the present study the partitioning data ($\log D$) in a PEG-2000/ $(\text{NH}_4)_2\text{SO}_4$ ABS, a diverse set of 29 organic molecules, were taken from the experimental work of Rogers et al.²⁰ The $\log D$ values range from 0.10 to 3.14 (Table 1).

The structures were drawn by using ISIS Draw 2.4 and preoptimized by using molecular mechanics force field method (MM+) available in Hyperchem 7.0.²⁹ Final optimizations were performed with MOPAC 7.0 (implemented in CODESSA PRO software) using the AM1 semiempirical method.³⁰ Constitutional, topological, geometrical, electrostatic, and semiempirical descriptors were calculated with CODESSA PRO software package.³¹ To select the “best” descriptor subset we used the BMLR (Best Multi Linear Regression) method. The BMLR is based on the (i) selection of the orthogonal descriptor pairs and (ii) extension of the correlation with the addition of new descriptors until the Fisher’s criteria becomes less than that of the best two-parameter correlation.

RESULTS AND DISCUSSION

In the current study, we present QSPR (Quantitative Structure–Property Relationship) models for $\log D$ (PEG/Salt), involving the theoretical descriptors, which have been calculated solely from molecular structure. To find the optimum number of descriptors describing $\log D$ for the current set of small organic probes, we analyzed multiparameter correlations containing up to seven descriptors. Figure 1 shows the relationships of R^2 and R^2_{cv} with the number of descriptors. As it can be seen in Figure 1, R^2 and R^2_{cv} rise steeply as the number of parameters increases from one to seven and the breakpoint occurs at the third descriptor. Therefore, we used the best correlation equation with three descriptors (eq 2) for the basic analysis. The plot of observed vs predicted $\log D$ values is illustrated in Figure 2.

$$\log D = (0.483 \pm 0.098) + (0.002 \pm 0.0001)G_b - (2.027 \pm 0.483)\text{SE}_{\text{OK}} + (2.062 \pm 0.267)q_{\text{min}} \quad (2)$$

$$N = 29; n = 3; R^2 = 0.967; R^2_{cv} = 0.956; F = 245.070; s^2 = 0.017$$

where N is the number of datapoints; n is the number of parameters; R^2 is the squared correlation coefficient; R^2_{cv}

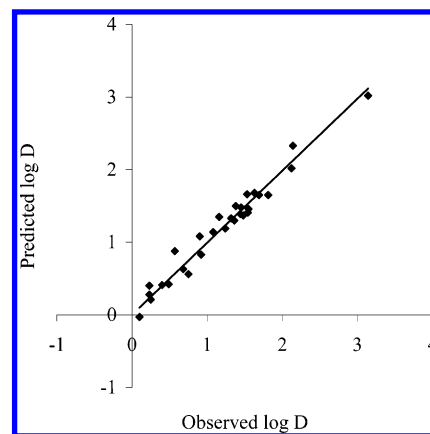


Figure 2. Plot of predicted vs observed values of $\log D$ using the three-parameter model (eq 2).

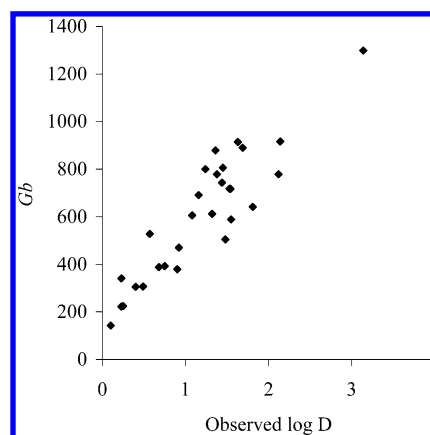


Figure 3. Distributions between $\log D$ and the G_b .

is the squared cross-validated correlation coefficient; F is the Fisher’s criterion; and s^2 is the squared standard error.

The first descriptor in the three-parameter model (eq 2), the gravitational index for all bonds (G_b), is defined by eq 3.

$$G_b = \sum_{i < j}^{N_b} \frac{m_i m_j}{r_{ij}^2} \quad (3)$$

In eq 3, m_i and m_j are the atomic masses of atoms i and j , r_{ij} is the interatomic distance between the atoms i and j , and N_b is the number of chemical bonds in the molecule.

G_b reflects preferential solubility in the PEG-rich phase of compounds of complex atomic composition. The G_b index may be regarded as responsible for the cavity-forming effects, because it accounts significantly for molecular size and the London dispersion forces.³² As it is clearly seen from Figure 3, $\log D$ correlates almost linearly with the gravitational index.

The second descriptor in eq 2 is the image of the Onsager–Kirkwood solvation energy (SE_{OK}). This descriptor is defined by eq 4, where μ is the dipole moment of the molecule calculated on the basis of AM1 Mulliken charges in our case and MW is the molecular weight.

$$\text{SE}_{\text{OK}} = \frac{\mu^2}{\text{MW}} \quad (4)$$

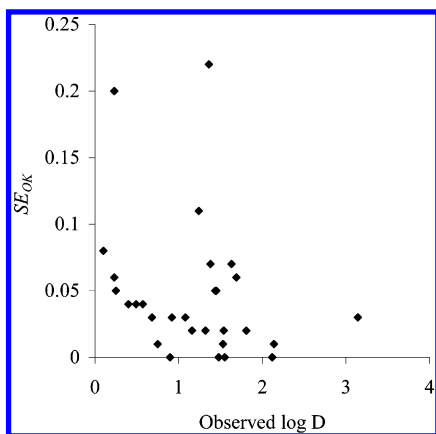


Figure 4. Distributions between $\log D$ and the SE_{OK} .

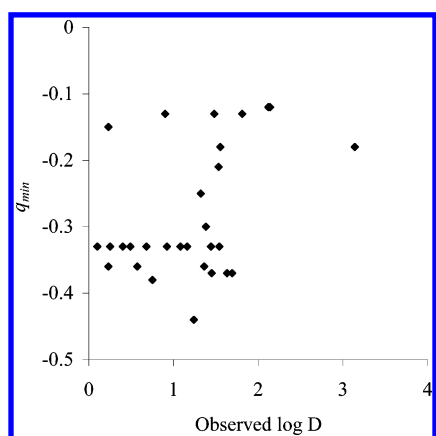


Figure 5. Distributions between $\log D$ and the q_{min} . Interrelation of the descriptors.

The descriptor SE_{OK} is a measure of the polarity of the solvent; its large negative coefficient indicates the reduced distribution of highly polar solutes into the polymer-rich phase. This descriptor is reminiscent of the Onsager reaction field, which sets the self-polarization of a solute molecule as proportional to its dipole moment.³³ Such polarization through the surrounding solvent medium contributes significantly to the short-range intermolecular interactions. In the case of ABS, it is reasonable to assume that the salt-rich phase, containing water dipoles and highly polarizable anions, causes a greater reaction field than the polymer-rich phase. These simple physical considerations suggest that the higher the SE_{OK} term the smaller the $\log D$ value. The scatter plot of $\log D$ values against SE_{OK} displayed in Figure 4 supports such a conclusion.

The minimal net atomic charge, q_{min} , presumably contributes both to the intensity of general electrostatic, in particular Coulombic, interactions and to hydrogen bonding (HB). The q_{min} descriptor can explicitly be related to hydrogen bonding, because a high positive value on hydrogen atoms implies good hydrogen bonding donor propensity, whereas a high positive value on heteroatoms (N, O, F) implies good acceptor ability of these atoms. The sign of q_{min} is positive that indicates to the higher basicity of the PEG-rich phase as compared to the salt-rich one. Although the distribution of the descriptor values is normal and the regression coefficient bears the positive sign, Figure 5 does not demonstrate any clear trend: additional analysis has revealed that for acidic species (HB donors) the correlation is rather

Table 2. Interrelation of the Descriptors (R^2 Values)

	G_b	SE_{OK}	q_{min}
G_b	1	0.0024	0.0055
SE_{OK}	0.0024	1	0.0819
q_{min}	0.0055	0.0819	1

Table 3. Verification of Statistical Validity of the Model

training sets	R^2	s^2	predicted sets	R^2	s^2
2 & 3	0.976	0.013	1	0.954	0.019
1 & 3	0.956	0.017	2	0.981	0.013
1 & 2	0.970	0.014	3	0.961	0.019
average	0.967	0.015	average	0.965	0.017

significant and negative ($R^2 = 0.85$), whereas for HB acceptors the correlation is less than 0.5 and positive.

We also studied possible intercorrelations between the descriptors involved in the three-parameter model (eq 2). As shown in Table 2, the descriptors are strongly orthogonal.

VALIDATION OF THE THREE-PARAMETER MODEL

To further demonstrate the absence of chance correlation, the whole data set was divided into three subsets (by using numbers 1, 4, 7, etc. 2, 5, 8, etc. and 3, 6, 9, etc.), and each subset was predicted by using the other two subsets as the training set. In this procedure, the same descriptors were retained in the correlation equation, but the coefficients were allowed to vary. The results are shown in Table 3, with average training quality of $R^2 = 0.967$ and average predicting quality of $R^2 = 0.965$, which indicates that the proposed model has a high statistical stability and validity.

BACKGROUND OF THE ONE- AND TWO-PARAMETER MODELS

For the sake of transparency and consistency, we also present one- and two-parameter models, derived for $\log D$ (Table 4). The best two-parameter model, calculated with CODESSA PRO (Table 4, eq 6), contains the same first two descriptors already involved in the three-parameter model— G_b and q_{min} .

We find that in the case of the single-parameter model (Table 4, eq 5), the so-called “best model” contains the Kier and Hall index (order 3) $^3X^v$ but not the gravitational index for all bonds. However, the Kier and Hall valence connectivity indices^{34–36} $^mX^v$ are defined by eq 5 and belong to the same “family” of descriptors, as does G_b .

$$^mX^v = \sum_{i=1}^{N_S} \prod_{k=1}^{m=1} \left(\frac{1}{\delta_k^v} \right)^{1/2} \quad (5)$$

In eq 5, δ_k^v is defined by eq 6.

$$\delta_k^v = \frac{(Z_k^v - H_k)}{(Z_k - Z_k^v - 1)} \quad (6)$$

In eqs 5 and 6, Z_k is the total number of electrons in the k th atom, Z_k^v is the number of valence electrons in the k th atom, H_k is the number of hydrogen atoms directly attached to the k th non-hydrogen atom, $m = 0$ is the atomic valence connectivity indices, $m = 1$ — one bond path valence connectivity indices, $m = 2$ — two bond fragment valence

Table 4. One- and Two-Parameter Models Derived for log *D*

eq no.	intercept	coef	desc	<i>t</i> -test	coef	desc	<i>t</i> -test	<i>R</i> ²	<i>R</i> ² _{CV}	<i>s</i> ²
Single-Parameter Models										
1	-0.239 ± 0.136	0.002 ± 0.0002	<i>G</i> _b	11.59				0.831	0.806	0.08
2	-0.561 ± 0.150	0.405 ± 0.033	⁰ χ ^v	12.39				0.851	0.834	0.07
3	-0.296 ± 0.153	0.633 ± 0.060	¹ χ ^v	10.49				0.803	0.774	0.10
4	0.219 ± 0.096	0.652 ± 0.055	² χ ^v	11.94				0.841	0.812	0.08
5	0.349 ± 0.070	1.026 ± 0.068	³ χ ^v	15.01				0.893	0.880	0.05
Two-Parameter Models										
6	0.471 ± 0.126	0.002 ± 0.0001	<i>G</i> _b	19.04	2.380 ± 0.328	<i>q</i> _{min}	7.26	0.944	0.927	0.03
7	0.014 ± 0.183	0.386 ± 0.026	⁰ χ ^v	14.67	1.743 ± 0.422	<i>q</i> _{min}	4.13	0.910	0.886	0.05
8	0.328 ± 0.195	0.605 ± 0.049	¹ χ ^v	12.46	1.974 ± 0.484	<i>q</i> _{min}	4.08	0.880	0.848	0.06
9	0.608 ± 0.190	0.619 ± 0.053	² χ ^v	11.78	1.207 ± 0.521	<i>q</i> _{min}	2.32	0.868	0.826	0.07
10	0.686 ± 0.149	0.980 ± 0.065	³ χ ^v	15.06	1.063 ± 0.422	<i>q</i> _{min}	2.52	0.914	0.889	0.04

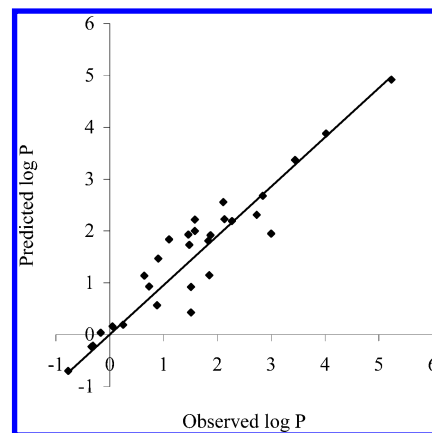
connectivity indices, and *m* = 3 three contiguous bond fragment valence connectivity indices, etc.

The Kier and Hall indices are related to molecular shape, as are the shadow descriptors. The involvement of the shape descriptors can be explained as the influence of steric factors on the probability of interaction with solvent. The descriptors are in accordance with the previous treatments of log *D* using Abraham's descriptors,^{37,38} which also reflects the strong contribution of molecular size. The significant positive coefficient of size related descriptors in the models indicate the importance of the cavity-forming forces in the PEG/salt distribution.

As mentioned, the gravitational index (*G*_b) and the Kier and Hall indices can be considered as belonging to the same group of descriptors. The *G*_b and the Kier and Hall indices are also highly intercorrelated, thus within the present data set *G*_b and ³χ^v are correlated as 0.86 in terms of the squared intercorrelation coefficient. Importantly, although the best single-parameter model is that containing the Kier and Hall index (order 3), this descriptor vanishes for six of the total 29 compounds (methanol, ethanol, 2-propanol, acetonitrile, acetic acid, and methyl iodide) out of the total of 29 (Table 1). Therefore we decided to study the single-parameter model with *G*_b (Table 4, eq 1) and also with the lower order (0, 1, and 2) Kier and Hall indices—⁰χ^v, ¹χ^v, and ²χ^v, respectively (Table 4, eqs 2–4). Additionally, we also derived two-parameter models where the gravitational index *G*_b replaces the Kier and Hall indices order 0, 1, 2, and 3 (Table 4, eqs 7–10). The data in Table 4 show that two-parameter models with Kier and Hall indices are less powerful than the corresponding two-parameter model with the gravitational index. For the sake of comparison, we also give the *R*² values for the three-parameter models with Kier and Hall indices (order 0, 1, 2, and 3), which are 0.910, 0.882, 0.871, and 0.918, respectively.

A comparison of all these models (one-, two-, and three-parameter) with the gravitational index or the Kier and Hall indices as major descriptors affirms the importance of the "shape descriptors" for describing the partition behavior of small organic compounds in aqueous biphasic systems. We have also shown that despite the best single-parameter model includes the Kier and Hall index (order 3), nevertheless, the best two- and three-parameter models utilize the gravitational index for all bonds.

It is fairly obvious that models presented above have some limitations because they have been calculated using log *D* values in a specific ABS (PEG-2000/(NH₄)₂SO₄). The salt and PEG may vary, and the current models may not work

**Figure 6.** Plot of predicted vs observed values of log *P* using the three-parameter model (eq 7).

properly for predicting log *D* in different ABS. Recently, Rogers et al.¹⁹ showed the importance of the TLL (tie line length) for describing the ABS. TLL represents the orthogonal sum of the difference in polymer (ΔPEG) and salt (Δsalt) concentrations between the phases and can be used for normalizing the log *D* scale (e.g. log *D*/TLL).

LOG *P* AS AN INDEPENDENT VARIABLE

For comparison, we also derived a three-parameter model (eq 7) for log *P* (octanol/water) with the same descriptors, which were used for log *D* in eq 2. Respective correlation of the experimental vs observed log *P* values is illustrated in Figure 6.

$$\log P = (0.901 \pm 0.364) + (0.004 \pm 0.0004)G_b - (6.249 \pm 1.788)SE_{OK} + (5.073 \pm 0.989)q_{\min} \quad (7)$$

$$N = 29; n = 3; R^2 = 0.878; R^2_{cv} = 0.837;$$

$$F = 64.68; s^2 = 0.236$$

Finally, a convincing correlation between observed and predicted log *D* was found, using calculated log *P* values (Table 1, eq 7) as an independent variable. The obtained single-parameter model is given in eq 8, and the corresponding correlation is illustrated in Figure 7.

$$\log D = (0.381 \pm 0.051) + (0.519 \pm 0.025) \log P_{(\text{pred})} \quad (8)$$

$$N = 29; n = 1; R^2 = 0.940; R^2_{cv} = 0.930;$$

$$F = 424.099; s^2 = 0.029$$

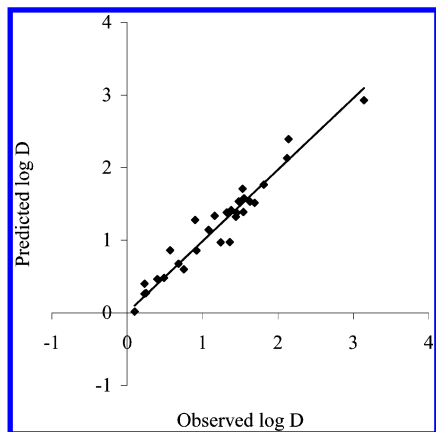


Figure 7. Plot of predicted vs observed values of log D using a single-parameter model (eq 8) with log $P_{(pred)}$.

Statistical parameters of eq 8 and Figure 7 once more affirm the strong correlation between log P and log D as was noted several years ago.^{8,9}

CONCLUSIONS

The partitioning of 29 small organic probes in a PEG-2000/(NH₄)₂SO₄ ABS was satisfactorily described with the three-parameter QSPR model (eq 2, $R^2 = 0.967$, $R^2_{cv} = 0.956$). All the descriptors involved were calculated solely from the chemical structures of compounds and have definite physical meaning corresponding to different intermolecular interactions. A single-parameter model involving calculated log P (octanol/water) values as an independent variable also demonstrates high statistical quality. The results described in this paper should help improve our understanding and prediction of partition coefficients in PEG/salt ABS for structurally diverse compounds.

ACKNOWLEDGMENT

The research at The Universities of Alabama and Florida was supported by the Division of Chemical Sciences, Office of Basic Energy Research, U.S. Department of Energy (Grant DE-FG02-96ER14673) and by the U.S. Environmental Protection Agency's STAR program through grant number R-82825701-0. (Although the research described in this article has been funded in part by EPA, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.) Additional support was provided to the Center for Green Manufacturing from the National Science Foundation Grant EPS-9977239. We would like to thank Dr. Scott K. Spear (Center for Green Manufacturing) for his assistance in the preparation of this manuscript.

REFERENCES AND NOTES

- (1) Albertsson, P.-A. Partition of Proteins in Liquid Polymer-Polymer Two-Phase Systems. *Nature* **1958**, *182*, 709–711.
- (2) Albertsson, P.-A. *Partition of Cell Particles and Macromolecules*; Wiley & Sons: New York, 1986.
- (3) Rogers, R. D.; Eiteman, M. A. *Aqueous Biphasic Separations: Biomolecules to Metal Ions*; Plenum Press: New York, 1995.
- (4) Zaslavsky, B. Y. *Aqueous Two-Phase Partitioning: Physical Chemistry and Bioanalytical Applications*; Marcel Dekker: New York, 1994.
- (5) Zaslavsky, B. Y.; Mestechkina, N. M.; Miheeva, L. M.; Rogozhin, S. V.; Bakalkin, G. Ya.; Rjzhzhsy, G. G.; Chetverina, E. V.; Asmuko, A. A.; Bespalova, J. D.; Korobov, N. V.; Chichenkov, O. N. Correlation of Hydrophobic Character of Opioid Peptides with their Biological Activity Measured in Various Bioassay Systems. *Biochem. Pharmacol.* **1982**, *31*, 3757–3762.

- (6) Diamond, A. D.; Hsu, J. T. Fundamental Studies of Biomolecule Partitioning in Aqueous Two-Phase Systems. *Biotechnol. Bioeng.* **1989**, *34*, 1000–1014.
- (7) Eiteman, M. A.; Gainer, J. L. Partition of Isomeric Dipeptides in Poly-(Ethylene Glycol)/Magnesium Sulfate Aqueous Two-Phase Systems. *Biochim. Biophys. Acta* **1991**, *1073*, 451–455.
- (8) Eiteman, M. A.; Gainer, J. L. A Model for the Prediction of Partition Coefficients in Aqueous Two-Phase Systems. *Bioseparation* **1991**, *2*, 31–41.
- (9) Eiteman, M. A.; Gainer, J. L. Predicting Partition Coefficients in Polyethylene Glycol-Potassium Phosphate Aqueous Two-Phase Systems. *J. Chromatogr.* **1991**, *586*, 341–346.
- (10) Eiteman, M. A.; Gainer, J. L. Peptide Hydrophobicity and Partitioning in Poly (Ethylene Glycol)/Magnesium Sulfate Aqueous Two-Phase Systems. *Biotechnol. Prog.* **1990**, *6*, 479–484.
- (11) Rogers, R. D.; Bond, A. H.; Bauer, C. B. Aqueous Biphasic Systems for Liquid/Liquid Extraction of f-Elements Utilizing Polyethylene Glycols. *Sep. Sci. Technol.* **1993**, *28*, 139–153.
- (12) Rogers, R. D.; Bond, A. H.; Baucer, C. B.; Zhang, J.; Griffin, S. T. Metal Ion Separations in Polyethylene Glycol-Based Aqueous Biphasic Systems: Correlation of Partitioning Behavior with Available Thermodynamic Hydration Data. *J. Chromatogr. B* **1996**, *680*, 221–229.
- (13) Rogers, R. D.; Bauer, C. B. Partitioning Behavior of Group 1 and 2 Cations in Poly(Ethylene Glycol)-Based Aqueous Biphasic Systems. *J. Chromatogr. B* **1996**, *680*, 237–241.
- (14) Rogers, R. D.; Griffin, S. T. Partitioning of Mercury in Aqueous Biphasic Systems and on ABEC Resins. *J. Chromatogr. B* **1998**, *711*, 277–283.
- (15) Rogers, R. D.; Willauer, H. D.; Griffin, S. T.; Huddleston, J. G. Partitioning of Small Organic Molecules in Aqueous Biphasic Systems. *J. Chromatogr. B* **1998**, *711*, 255–263.
- (16) Willauer, H. D.; Huddleston, J. G.; Griffin, S. T.; Rogers, R. D. Partitioning of Aromatic Molecules in Aqueous Biphasic Systems. *Sep. Sci. Technol.* **1999**, *34*, 1069–1090.
- (17) Li, M.; Willauer, H. D.; Huddleston, J. G.; Rogers, R. D. Temperature Effects on Polymer-Based Aqueous Biphasic Extraction Technology in the Paper Pulp Process. *Sep. Sci. Technol.* **2001**, *36*, 835–847.
- (18) Willauer, H. D.; Huddleston, J. G.; Li, M.; Rogers, R. D. Investigation of Aqueous Biphasic Systems for the Separation of Lignins from Cellulose in the Paper Pulp Process. *J. Chromatogr. B* **2000**, *743*, 127–135.
- (19) Willauer, H. D.; Huddleston, J. G.; Rogers, R. D. Solute Partitioning in Aqueous Biphasic Systems Composed of Polyethylene Glycol and Salt: The Partitioning of Small Neutral Organic Species. *Ind. Eng. Chem. Res.* **2002**, *41*, 1892–1904.
- (20) Willauer, H. D.; Huddleston, J. G.; Rogers, R. D. Solvent Properties of Aqueous Biphasic Systems Composed of Polyethylene Glycol and Salt Characterized by the Free Energy of Transfer of a Methylene Group Between the Phases and by a Linear Solvation Energy Relationship. *Ind. Eng. Chem. Res.* **2002**, *41*, 2591–2601.
- (21) Huddleston, J. G.; Willauer, H. D.; Boaz, K. R.; Rogers, R. D. Separation and Recovery of Food Coloring Dyes Using Aqueous Biphasic Extraction Chromatographic Resins. *J. Chromatogr. B* **1998**, *711*, 237–244.
- (22) Huddleston, J. G.; Ingenito, C. C.; Rogers, R. D. Partitioning Behavior of Porphyrin Dyes in Aqueous Biphasic Systems. *Sep. Sci. Tech.* **1999**, *34*, 1091–1101.
- (23) Huddleston, J. G.; Willauer, H. D.; Rogers, R. D. Solvatochromic Studies in Polyethylene Glycol-Salt Aqueous Biphasic Systems. *J. Chromatogr. B* **2000**, *743*, 137–149.
- (24) Huddleston, J. G.; Willauer, H. D.; Rogers, R. D. The Solvatochromic Properties, α , β , and π^* , of PEG-salt Aqueous Biphasic Systems. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4065–4070.
- (25) Huddleston, J. G.; Willauer, H. D.; Griffin, S. T.; Rogers, R. D. Aqueous Polymeric Solutions as Environmentally Benign Liquid/Liquid Extraction Media. *Ind. Eng. Chem. Res.* **1999**, *38*, 2523–2539.
- (26) Abraham, M. H.; Zissimos, A. M.; Huddleston, J. G.; Willauer, H. D.; Rogers, R. D.; Acree, W. E., Jr. Some Novel Liquid Partitioning Systems: Water-Ionic Liquids and Aqueous Biphasic Systems. *Ind. Eng. Chem. Res.* **2003**, *42*, 413–418.
- (27) Sangster, J. *Octanol-Water Partition Coefficients: Fundamentals and Physical Chemistry*; John Wiley and Sons: New York, 1997; pp 113–127.
- (28) Katritzky, A. R.; Wang, Y.; Slid, 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.
- (29) <http://www.hyper.com>.
- (30) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (31) <http://www.codessa-pro.com>.

- (32) Katritzky, A. R.; Mu, L.; Lobanov, V. S.; Karelson, M. Correlation of Boiling Points With Molecular Structures. 1. A Training Set of 298 Diverse Organics and a Test Set of 9 Simple Inorganics. *J. Phys. Chem.* **1996**, *100*, 10400–10407.
- (33) Onsager, L. Electric Moments of Molecules in Liquids. *J. Am. Chem. Soc.* **1936**, *58*, 1486–1493.
- (34) Kier, L. B.; Hall, L. H. The Nature of Structure–Activity Relationships and their Relation to Molecular Connectivity. *Eur. J. Med. Chem.* **1977**, *12*, 307–312.
- (35) Kier, L. B.; Hall, L. H. Derivation and Significance of Valence Molecular Connectivity. *J. Pharm. Sci.* **1981**, *70*, 583–589.
- (36) Karelson, M. *Molecular Descriptors in QSAR/QSPR*; J. Wiley & Sons: New York, 2000.
- (37) Abraham, M. H. Scales of Solute Hydrogen-Bonding: Their Construction and Application to Physicochemical and Biochemical Processes. *Chem. Soc. Rev.* **1993**, *22*, 73–83.
- (38) Platts, J. A.; Abraham, M. H.; Butina, D.; Hersey, A. Estimation of Molecular Linear Free Energy Relationship Descriptors by a Group Contribution Approach. 2. Prediction of Partition Coefficients. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 71–80.

CI034194O