A Quantitative Structure−Property Relationship Analysis of logP for Disubstituted Benzenes

Jian-Wei Zou,* Wen-Na Zhao, Zhi-Cai Shang, Mei-Lan Huang, Ming Guo, and Qing-Sen Yu

Department of Chemistry, Zhejiang University, Hangzhou, 310027, People's Republic of China

Received: April 21, 2002; In Final Form: September 24, 2002

Ab initio HF/6-31G* optimization and electrostatic potentials have been computed for a number of disubstituted benzenes. A good six parameter correlation between logP and the computed descriptors for a training set of 103 representative disubstituted benzenes is presented. The predictive power of this model has been demonstrated on a test set of 111 disubstituted benzenes and several polysubstituted benzenes. Of these descriptors introduced, ΣV_S^-, a new descriptor proposed in the present work and defined as the sum of the surface minima values of the electrostatic potential, gives the most significant contribution to the logP. This descriptor, together with the molecular volume V, is thought to be mainly responsible for variations in logP with the substituent groups. The influence of group−group interactions on the logP values of isomers has been investigated by treating the ortho-disubstituted benzenes separated from the others. It shows that the influences of interactions between two neighboring substituent groups are basically thought to be relevant to hydrogen bonding and are represented mainly by variations of the V_{min}, $V_{\text{s,max}}$, and polar surface area values of isomers. The applicability of the theoretical linear solvation energy relationship approach and the one proposed by Bodor have been tested and discussed.

1. Introduction

The logarithm of the partition coefficient between *n*-octanol and water (logP) is a frequently used parameter in many quantitative structure−activity relationship (QSAR) studies that have been developed in pharmaceutical, environmental, biochemical, and toxicological sciences.[1] Although logP is generally easy to determine, the accurate measurement is still time-consuming and expensive. Moreover, several studies are often concerned with molecules that are not yet synthesized and for which the logP is not known. Hence, it is of great significance for developing rapid and efficient methods to estimate or predict the logP value in advance.

At present, many methods of calculating logP have been proposed since the pioneering work[2] of Hansch and Fujita appeared about four decades ago, and several excellent reviews on this subject have been published.[1,3−6] Generally speaking, most methods can be divided into three classes based on (i) fragment contributions, (ii) atomic contributions, or (iii) molecular descriptors. Although both the fragment additive and the atom additive approaches can give reliable predictions, they cannot estimate logP for those molecules, which contain an unknown fragment or atom type. Many correction factors have to be added to the practical calculation due to the existence of intramolecular group−group interactions; moreover, the nature of these correction factors still remains unclear. It seems that the approaches based on molecular descriptors can avoid, in principle, such problems, because they treat molecules as a whole rather than a simple sum of parts, and the influences of group−group interactions are automatically taken into account in the determination or calculation of molecular descriptors. Nevertheless, it is of equal importance for physical organic chemists to understand how intramolecular group−group interactions influence the partition of solutes.

There are many group−group interactions involved and experimental partition coefficients available for disubstituted benzenes (compounds of the type XC_6H_4Y). These compounds are therefore very suited to be used to explore the physicochemical reason for variations in logP with substituted groups and their relative positions. Fujita[7] proposed a procedure to analyze and predict the logP values of these compounds by using so-called "bidirectional" Hammett type relationships. Similar works have been reported by Brandstrom[8] and Leo[9] almost at the same time. More recently, Brandstrom[10] presented their revised procedure. Although cross terms such as $\sigma_{(X)}\rho_{(Y)}$ were introduced in these studies to express the group−group interactions, these procedures often fail to apply for ortho-disubstituted benzenes, and the nature of the cross terms had not been disclosed. Camilleri et al.[11] have ever predicted the logP for numerous substituted benzenes by calculating the solvent accessible surface area of 12 previously defined component atoms and groups. Despite the fact that there is no need for a correction factor, this approach falls into the category of empirical fragment additive method, and it is still difficult to know the intrinsic reason for variations in logP with molecular structure. The approaches based on molecular descriptors have an advantage of elucidating the influences of group−group interactions on logP over the others, but so far, no such studies designed for this purpose have been found.

Bearing this in mind, we set out to analyze and predict the logP of disubstituted benzenes by correlating it with theoretical descriptors with explicit physical meaning. Especially, a set of statistically based indices devised by Politzer et al. using molecular electrostatic potential, which has been proven to be fairly effective for correlating and predicting the logP[12−14] as well as other physicochemical properties that reflect solvent−solute interactions,[15−20] will be used in the present study. The

---

* To whom correspondence should be addressed. Fax: +86-571-87951895. E-mail: jwzou@css.zju.edu.cn.

10.1021/jp025984o CCC: $22.00

TABLE 1: Compounds with Computed Descriptors Together with the Experimental and Predicted logP Using Eq 1[a]

		ΣV_S^-	V_{min}	$V_{s,max}$	V	PSA	μ^2/V	logP (exp)	logP (calcd)	residual
1	3-chloro-1-nitrobenzene	−18.3448	−1.9342	1.3476	118.86	35.91	0.1565	2.46	2.67	−0.21
2	4-chloro-1-nitrobenzene	−19.3872	−1.9370	1.2640	118.71	36.16	0.0833	2.39	2.32	0.07
3	2-chloro-1-nitrobenzen	−26.7425	−2.0058	1.3224	118.71	36.21	0.2627	2.24	1.88	0.36
4	1,3-dichlorobenzene	−11.8643	−0.5708	1.0830	110.66	0.00	0.0371	3.53	3.63	−0.10
5	1,4-dichlorobenzene	−11.7948	−0.5723	1.0322	110.48	0.00	0.0000	3.44	3.52	−0.08
6	1,2-dichlorobenzene	−17.9780	−0.9121	1.0459	110.18	0.00	0.0908	3.43	2.95	0.48
7	3-chlorophenol	−18.4588	−1.7169	2.7278	104.32	22.05	0.0086	2.50	2.44	0.06
8	4-cholrophenol	−19.9982	−1.7286	2.7562	105.00	22.05	0.0581	2.39	2.41	−0.02
9	2-cholrophenol	−17.5567	−1.7516	1.7401	104.18	20.92	0.0153	2.15	2.26	−0.11
10	3-nitrophenol	−23.3348	−2.0409	2.9530	113.18	57.99	0.1234	2.00	1.80	0.20
11	4-nitrophenol	−27.6719	−2.2010	3.0580	113.08	58.11	0.2614	1.91	1.62	0.29
12	2-nitrophenol	−21.6796	−1.8991	1.5853	111.90	52.47	0.1415	1.79	1.69	0.10
13	3-chloroaniline	−25.2424	−1.0354	1.9196	107.62	26.86	0.1051	1.88	1.72	0.16
14	2-chloroaniline	−20.6331	−1.0927	1.9125	107.87	25.94	0.0401	1.90	2.14	−0.24
15	4-chloroaniline	−23.9486	−1.0449	1.9156	107.84	27.10	0.0811	1.88	1.82	0.06
16	3-chlorobenzene sulfonamide	−30.9519	−2.1563	2.0952	139.78	68.45	0.1152	1.29	1.44	−0.15
17	2-chlorobenzene sulfonamide	−34.6339	−2.2912	2.0355	140.88	68.22	0.1291	0.74	1.04	−0.30
18	4-chlorobenzene sulfonamide	−26.4059	−2.1487	2.0980	140.76	68.09	0.0791	1.24	1.95	−0.71
19	3-nitroaniline	−29.9248	−2.2194	2.1636	116.49	63.07	0.3196	1.37	1.25	0.12
20	2-nitroaniline	−26.0096	−2.2485	2.1749	115.17	57.81	0.2216	1.85	1.53	0.32
21	4-nitroaniline	−31.8367	−2.4354	2.3005	116.16	62.96	0.4830	1.39	1.43	−0.04
22	3-nitrobenzene sulfonamide	−35.8184	−1.9724	2.3348	149.21	104.41	0.1849	0.55	0.78	−0.23
23	2-nitrobenzene sulfonamide	−42.5904	−2.3198	2.3097	148.27	98.90	0.4018	0.34	0.52	−0.18
24	4-nitrobenzene sulfonamide	−34.9468	−1.8987	2.3656	149.43	104.36	0.1166	0.64	0.75	−0.11
25	1,3-benzenediol	−27.7731	−1.9593	2.5015	98.68	44.14	0.0213	0.80	0.56	0.24
26	1,2-benzenediol	−25.6862	−2.1520	2.7056	98.38	41.83	0.0722	0.88	1.02	−0.14
27	1,4-benzenediol	−28.3135	−2.0331	2.4946	98.66	44.35	0.0000	0.59	0.43	0.16
28	3-aminophenol	−32.1449	−2.1273	2.4326	102.11	48.76	0.0920	0.21	0.21	0.00
29	2-aminophenol	−27.6327	−1.7599	2.2555	101.75	46.94	0.0045	0.62	0.55	0.07
30	4-aminophenol	−30.3378	−2.2535	2.3118	101.61	49.19	0.0514	0.04	0.25	−0.21
31	1,3-diaminobenzene	−32.5867	−1.4585	1.5941	104.82	54.05	0.0237	−0.33	−0.19	−0.14
32	1,2-diaminobenzene	−30.7995	−1.5132	2.2649	105.09	51.21	0.0558	0.15	0.38	−0.23
33	1,4-diaminobenzene	−32.0807	−1.6918	1.5296	105.19	54.21	0.0000	−0.30	−0.23	−0.07
34	3-trifluoromethyl-1-nitrobenzene	−23.3010	−1.8881	1.3815	135.90	34.41	0.1284	2.62	2.73	−0.11
35	2-trifluoromethyl-1-nitrobenzene	−28.1838	−1.9445	1.4066	135.13	34.88	0.2775	2.58	2.47	0.11
36	4-trifluoromethyl-1-nitrobenzene	−21.6296	−1.8481	1.3143	135.39	35.95	0.0381	2.55	2.64	−0.09
37	3-cyano-1-nitrobenzene	−30.6714	−2.1219	1.5276	122.11	54.48	0.1818	1.17	1.03	0.14
38	2-cyano-1-nitrobenzene	−35.0055	−2.3286	1.5615	121.61	52.93	0.4639	1.02	1.20	−0.18
39	4-cyano-1-nitrobenzene	−27.7511	−2.0371	1.4684	122.23	54.48	0.0001	1.19	0.93	0.26
40	3-trifluoromethylphenol	−21.2381	−1.6684	2.7621	121.24	22.05	0.0214	2.95	2.83	0.12
41	2-trifluoromethylphenol	−16.7775	−1.6623	2.1825	120.94	19.65	0.0275	2.80	3.24	−0.44
42	4-trifluoromethylphenol	−22.0040	−1.6049	2.8492	121.01	22.01	0.0856	2.82	2.92	−0.10
43	3-cyanophenol	−25.7278	−2.4257	2.8194	108.16	40.69	0.1146	1.70	1.54	0.16
44	2-cyanophenol	−21.6549	−2.2405	2.1960	107.35	39.25	0.1295	1.61	1.90	−0.29
45	4-cyanophenol	−26.2537	−2.5806	2.9633	107.65	40.53	0.2462	1.60	1.81	−0.21
46	3-nitrobenzoic acid	−29.7864	−2.0137	2.7705	132.14	74.17	0.2637	1.83	1.76	0.07
47	2-nitrobenzoic acid	−34.9137	−2.0363	2.5641	131.85	74.00	0.1616	1.46	0.80	0.66
48	4-nitrobenzoic acid	−27.6540	−1.9233	2.8470	132.47	74.00	0.1220	1.89	1.71	0.18
49	2-trifluoromethylaniline	−21.8218	−0.9601	1.9745	123.91	24.39	0.0671	2.41	2.77	−0.36
50	3-trifluoromethylaniline	−26.1625	−0.9864	1.9623	123.94	27.06	0.1304	2.29	2.34	−0.05
51	4-trifluoromethylaniline	−27.5827	−1.0033	2.0397	124.63	27.02	0.1902	2.39	2.37	0.02
52	3-hydroxylbenzaldehyde	−25.4132	−2.3148	2.6094	109.74	39.47	0.0386	1.38	1.42	−0.04
53	2-hydroxylbenzaldehyde	−21.9632	−2.1568	1.2153	108.68	34.16	0.1069	1.81	1.66	0.15
54	4-hydroxylbenzaldehyde	−28.0933	−2.5125	2.8442	109.75	39.53	0.1894	1.35	1.52	−0.17
55	3-chlorotoluene	−17.1683	−0.8155	0.8213	113.45	0.00	0.0501	3.28	3.02	0.26
56	2-chlorotoluene	−16.5426	−0.8126	0.8107	113.04	0.00	0.0360	3.42	3.04	0.38
57	4-chlorotoluene	−17.2721	−0.8277	0.8219	113.46	0.00	0.0553	3.33	3.02	0.31
58	4-chloroanisole	−21.1567	−1.6685	1.1061	121.19	10.10	0.0712	2.78	2.68	0.10
59	2-chloroanisole	−17.6662	−1.6723	0.7847	121.75	9.39	0.0031	2.68	2.87	−0.19
60	3-chloroanisole	−19.1148	−1.6686	0.9775	121.48	10.01	0.0204	2.98	2.77	0.21
61	2-aminobenzoic acid	−27.4698	−2.1260	2.3555	120.11	59.86	0.0119	1.21	1.06	0.15
62	4-aminobenzoic acid	−33.4597	−2.5244	2.1594	120.46	64.99	0.1432	0.83	0.46	0.37
63	3-nitrotoluene	−24.0586	−2.1624	1.1439	121.60	36.04	0.2333	2.42	2.18	0.24
64	2-nitrotoluene	−21.5600	−2.1655	1.1146	121.19	34.12	0.1989	2.30	2.41	−0.11
65	4-nitrotoluene	−24.2097	−2.1911	1.0877	121.15	36.11	0.2599	2.37	2.19	0.18
66	3-nitroanisole	−23.9608	−2.1009	1.0751	130.19	46.01	0.1381	2.16	2.09	0.07
67	2-nitroanisole	−28.6592	−2.1694	1.1744	129.93	44.74	0.2123	1.73	1.74	−0.01
68	4-nitroanisole	−28.0681	−2.2499	1.3767	129.51	46.16	0.2912	2.03	2.02	0.01
69	3-methylphenol	−22.6721	−2.0329	2.4892	107.87	22.05	0.0280	1.96	2.00	−0.04
70	2-methylphenol	−22.0122	−1.8962	2.4384	107.43	21.07	0.0132	1.95	2.05	−0.10
71	4-methylphenol	−21.8408	−2.0394	2.4642	107.31	22.08	0.0198	1.94	2.05	−0.11
72	3-methylaniline	−22.5554	−1.3641	1.6987	110.68	26.98	0.0192	1.40	1.84	−0.44
73	2-methylaniline	−23.7723	−1.3960	1.7308	110.41	25.39	0.0229	1.32	1.73	−0.41

TABLE 1 (Continued)

		ΣV_S^-	V_{min}	$V_{s,max}$	V	PSA	μ^2/V	logP (exp)	logP (calcd)	residual
74	4-methylaniline	−21.6059	−1.4032	1.6923	110.85	26.96	0.0179	1.39	1.95	−0.56
75	3-methoxyaniline	−29.3734	−2.6868	1.5426	118.86	36.99	0.0156	0.93	0.92	0.01
76	2-methoxyaniline	−25.6381	−2.6689	1.5357	118.99	35.96	0.0039	1.18	1.37	−0.19
77	4-methoxyaniline	−27.6262	−2.9124	1.5087	119.23	37.30	0.0305	0.95	1.14	−0.19
78	1,3-dicyanobenzene	−30.7476	−2.1584	1.4894	117.18	37.10	0.1821	0.80	1.15	−0.35
79	1,4-dicyanobenzene	−28.2610	−2.1092	1.4264	117.10	37.07	0.0000	0.93	0.98	−0.05
80	1,2-dicyanobenzene	−35.4810	−2.2499	1.4888	116.32	37.02	0.4991	0.99	1.32	−0.33
81	1,3-phthalic acid	−31.8443	−2.1827	2.5403	136.66	76.44	0.1157	1.66	1.18	0.48
82	1,2-phthalic acid	−34.4792	−3.6722	2.3629	135.79	71.35	0.1131	0.73	0.68	0.05
83	1,4-phthalic acid	−27.4369	−2.0036	2.6229	136.15	76.15	0.0648	2.00	1.62	0.38
84	3-methylbenzoic acid	−25.3199	−2.2588	2.4165	126.45	38.05	0.0361	2.37	2.08	0.29
85	2-methylbenzoic acid	−23.1940	−2.0998	2.4019	125.63	35.84	0.0271	2.46	2.34	0.12
86	4-methylbenzoic acid	−24.4004	−2.2975	2.2540	125.99	38.25	0.0492	2.27	2.15	0.12
87	3-methoxybenzoic acid	−27.7337	−2.1347	2.4216	134.31	47.92	0.0041	2.02	1.84	0.18
88	2-methoxybenzoic acid	−35.5965	−3.0436	1.5608	133.04	42.31	0.3851	1.59	1.52	0.07
89	4-methoxybenzoic acid	−28.4645	−2.3312	2.3070	134.30	48.12	0.0403	1.96	1.78	0.18
90	*m*-xylene	−14.4087	−0.9899	0.5654	115.87	0.00	0.0004	3.20	3.22	−0.02
91	*o*-xylene	−15.9699	−1.0002	0.5609	115.54	0.00	0.0017	3.12	3.02	0.10
92	*p*-xylene	−14.0327	−1.9815	0.5520	116.10	0.00	0.0000	3.15	3.15	0.00
93	3-methylanisole	−22.2681	−1.9669	0.7963	124.29	10.08	0.0094	2.66	2.38	0.28
94	2-methylanisole	−19.8425	−1.7647	0.8286	124.22	8.97	0.0116	2.74	2.74	0.00
95	4-methylanisole	−20.6569	−1.9715	0.8197	124.44	10.02	0.0126	2.66	2.60	0.06
96	3-nitroacetophenone	−27.1348	−2.0874	1.2048	140.15	52.82	0.0271	1.42	1.74	−0.32
97	2-nitroacetophenone	−32.0805	−2.2891	1.3016	140.57	51.61	0.2110	1.28	1.64	−0.36
98	4-nitroacetophenone	−28.4871	−2.0141	1.5179	140.47	52.74	0.1200	1.53	1.93	−0.40
99	3-nitrobenzoic acid methyl ester	−33.0856	−2.0648	1.2529	148.49	61.38	0.2817	1.89	1.83	0.06
100	2-nitrobenzoic acid methyl ester	−34.4039	−2.1510	1.2460	148.72	61.13	0.1143	1.66	1.25	0.41
101	3-chlorobenzoic acid	−19.6095	−2.0025	2.5440	123.17	38.13	0.1123	2.68	2.90	−0.22
102	2-chlorobenzoic acid	−24.4343	−2.5059	2.4466	123.08	36.96	0.0018	2.05	1.97	0.08
103	4-chlorobenzoic acid	−20.2842	−2.0217	2.4841	123.52	38.19	0.0204	2.65	2.58	0.07

[a] In units of ev for $V_{s,max}$, V_{min}, and ΣV_S^-, Å^3 for V, Å^2 for PSA, and Debye^2•Å^{-3} for μ^2/V.

purpose of this effort is (i) to test how well the theoretical descriptors correlate and predict the logP for disubstituted benzenes and (ii) to gain, through the correlation, an insight into the reason for variations in logP with chemical structures.

2. Methods

Molecular geometries were first optimized at the HF/6-31G* level by using the Gaussian 98w software package.[21] Using these molecular structures, many molecular descriptors including electrostatic potentials have been calculated. The 6-31G* basis set has been demonstrated to give accurate geometries[22] as well as electrostatic potentials. The electrostatic potentials were calculated with the grid control option set to "cube = 100". Then, the statistically based indices derived from these electrostatic potentials on ca. 100^3 points, mostly according to Politzer et al.,[17] were obtained. Molecular size descriptors such as molecular volume and surface area were calculated by using the modified Bodor method.[23] A correlation of combinations of the descriptors was established by multiple linear regression analysis.

3. Results and Discussion

The experimentally determined logP values were taken mainly from the compilation of Hansch and Leo,[24] with a few additions from Carl' one.[25] A group of 103 representative disubstituted benzenes (XC_6H_4Y, X and Y = Cl, OH, NH_2, NO_2, CF_3, CN, CHO, COOH, CH_3, OCH_3, $COCH_3$, and $COOCH_3$) was selected as the training set. Table 1 presents these compounds with their molecular descriptors. Also listed are their experimental and predicted logP values as well as the residues.

By using stepwise linear regression, a number of different models were tested and the best one on the basis of the present parameter set at a 95% confidence level was given as

$$\log P = 0.12194\ \Sigma V_S^- + 0.31513\ V_{s,max} + 0.04050\ V + 2.49876\ \mu^2/V - 0.01993\ \mathrm{PSA} + 0.12320\ V_{min} + 0.23163$$

$$n = 103,\ \mathrm{sd} = 0.247,\ r = 0.962,\ r_{cv} = 0.957,\ \text{and}\ F = 195.98 \quad (1)$$

where, throughout this paper, n is the number of data points submitted to the regression, sd is the standard deviation, r and r_{cv} are the correlation coefficient and the "leave-one-out" cross-validated correlation coefficient, respectively, and F is the overall statistical significance of the equation.

The correlation given by eq 1 contains six regression parameters. Three of them, V_{min}, $V_{s,max}$, and ΣV_S^-, are derived from molecular electrostatic potentials. V_{min} and $V_{s,max}$ are defined by Politzer et al. as the spatial minima and surface maxima of the electrostatic potential, respectively. The former measures the hydrogen bond-accepting tendency or hydrogen bond basicity of a molecule, whereas the latter measures the hydrogen bond-donating tendency or hydrogen bond acidity of a molecule.[16,17] We found that the V_{min} gave a slightly better correlation than the $V_{s,min}$, a quite similar quantity just defined as the surface minima of the molecular electrostatic potential.

In eq 1, V_{min} gives small but statistically significant contributions at the 96.3% confidence level by Student's *t*-test to logP. The sign of this term is positive, which means that the increase of hydrogen bond basicity of a solute would favor its partition in water. The same conclusion can be found in many other literatures.[26−29] Several researchers[12,27,29,30] have shown that the hydrogen bond acidity makes no significant contribution to the logP for their data set; however, the $V_{s,max}$ term in eq 1 reveals

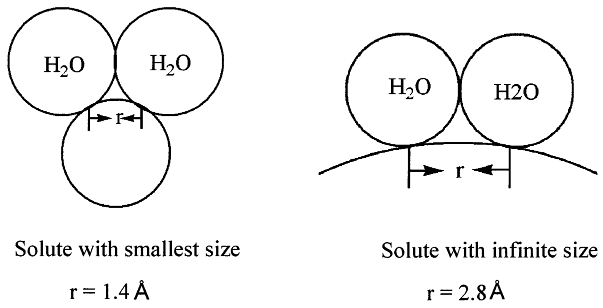

Figure 1. Schematic diagram of the cutoff value of the distance between two minimum electrostatic potentials introduced in ΣV_S^-.

that the hydrogen bond acidity is statistically significant and gives a positive contribution to logP for disubstituted benzenes. This finding is in accordance with the one of Kamlet et al.[28] who correlated the logP for a data set of 245 organic compounds including strong hydrogen bond donor solutes, as well as a subset containing 174 aromatic compounds, by using their linear solvation energy relationships model.

ΣV_S^- is defined as the sum of the surface minima values of the electrostatic potential. The sum started from the most negative potential on the molecular surface, and if two $V_{s,\min}$ appeared within 2.1 Å of each other, only the minimum with the most negative potential was included in the sum. This descriptor is similar to $\Sigma V_{\min}$ proposed previously by Haeberlein et al.[12] Differing from ΣV_S^-, $\Sigma V_{\min}$ is the sum of the minimum spatial electrostatic potential and the $V_{\min}$ values less negative than -147 kJ·mol^{-1} would not be included. According to Haeberlein et al., $\Sigma V_{\min}$ is also an effective hydrogen bond basicity descriptor, and the use of an empirical cutoff value (-147 kJ·mol^{-1}) is needed since very weak hydrogen bond interactions are taken into account in dipolarity/polarizability descriptors. For this reason, ΣV_S^- should not be viewed as a hydrogen bond basicity descriptor but one reflecting nonspecific intermolecular interactions despite the fact that similarities exist between each other. In fact, we found that ΣV_S^- correlated better to Π ($r = 0.83$), another descriptor derived from electrostatic potential, which has been proven to be a good descriptor reflecting nonspecific intermolecular interactions,[18,31] than to $V_{\min}$ or $V_{s,\min}$ (the correlation coefficients are 0.58 and 0.57, respectively).

It is noteworthy that the distance between two minimum electrostatic potentials introduced in both ΣV_S^- and $\Sigma V_{\min}$, 2.1 Å, is also an empirical cutoff value. Haeberlein et al. gave no interpretation for this cutoff value. In the present work, with an attempt to seek the best cutoff value, we changed the value with a step of 0.3 Å and calculated the ΣV_S^-. However, no improved correlation was found. Therefore, this cutoff value should be a reasonable one. In fact, if both solvent and solute molecules are treated simply as rigid spheres, and the solvent spheres can arrange close on the surface of a solute molecule to one another, the cutoff value of 2.1 Å is just the arithmetic mean value between the van der Waals radius of a water molecule (1.4 Å) and the diameter of a water molecule (2.8 Å). The value of 1.4 Å can also be viewed as the distance between two neighboring points of a water sphere in contact with a solute sphere with the smallest size (water is regarded as the smallest molecule except gas), while the value 2.8 Å can be viewed as the distance between two neighboring points of a water sphere in contact with a solute sphere with infinite size. The schematic diagram of the cutoff value can be seen in Figure 1.

PSA (polar surface area), which is defined as the van der Waals surface area occupied by the oxygen and nitrogen atoms as well as hydrogen atoms attached to them, is another descriptor introduced in eq 1. This quantity has been used successfully to correlate and predict several transport properties of drugs.[32−34] Although the so-called "dynamic" PSA (PSA$_d$),[35,36] a Boltzmann-weighted average value computed from all conformers within 2.5 kcal·mol^{-1} of the lowest energy conformer found during a detailed conformational search, has been demonstrated to be a better descriptor of molecular surface properties than the static PSA that utilizes only a single conformer, it takes much more computational cost, and a strong linear correlation has been found[33,34] between each other. By definition, PSA is often viewed as a descriptor for reflecting the hydrogen bond-forming capacity of a solute.[37,38] This is easy to understand since a large PSA value means that the corresponding solute can provide more hydrogen bond acceptor sites and/or hydrogen bond donor sites. In one recent study, Osterberg et al.[39] found that the PSA, in some cases, can be simplified by using the number of hydrogen bond-forming atoms. In eq 1, this descriptor gives negative contribution to logP. This is quite reasonable if one has noted that water is either a better hydrogen bond donor or a better hydrogen bond acceptor than octanol, and the increase of PSA would be naturally in favor of the partition in the water where stronger hydrogen bonding can be formed.

Also included in eq 1 are molecular volume (V) and μ^2/V. Molecular volume is often viewed as a cavity term. Some researchers[12,40] have shown that the surface area gave slightly better correlation than the volume. However, a less satisfactory correlation was found in the present work when the surface area was substituted for the volume. This is consistent with the result reported recently by Bodor et al.[29] The sign of the volume is positive, which indicates that the solute with a larger size would tend to distribute into octanol. This makes sense since the cavity term is a measure of the energy needed to overcome the cohesive forces in order to form a cavity for the solute, and the larger molecules would tend to be excluded from the more polar solvent water. The μ^2/V term, which was also presented in a previous study,[12] has its origin in the reaction field theory of Kirkwood and Onsager.[41] Mu et al.[42] have used a similar descriptor μ/V (dipolar density) to correlate S′ parameters of the unified solvent polarity scale. They found that this descriptor could give better correlation than the Kirkwood−Onsager solvation energy. However, the substitution of the μ/V for μ^2/V descriptor would lower the correlation slightly ($r = 0.959$, sd $= 0.254$). It should be noted that our μ^2/V or μ/V term gives a positive contribution to logP, and the interpretation of this behavior, as Haeberlein et al.[12] have pointed out, is therefore difficult or not straightforward.

As there is more than one hydrogen-bonding term ($V_{\min}$, $V_{s,\max}$, and PSA) and polarity term (ΣV_S^- and μ^2/V) presented in eq 1, it is necessary to examine the stability of our regression. Upon investigating the collinearity of variables in eq 1, we obtain the variance inflation factor (VIF) for each descriptor, and they are ΣV_S^- (4.02), $V_{s,\max}$ (1.82), $V_{\min}$ (1.58), V (1.75), PSA (5.07), and μ^2/V (1.45), respectively. The VIF values for ΣV_S^- and PSA are, although somewhat large, considered acceptable (according to statistics principle,[43] a value of 1.0 is indicative of no correlation, while a value of under 10.0 is statistically satisfactory).

To test the predictive power of the model given by eq 1, we first performed leave-one-out cross-validated analysis on the training set; the result was satisfactory (the r_{cv} reached to 0.957). A plot of predicted vs experimental logP values is shown in Figure 2. Then, a group of 111 disubstituted benzenes was used as a test set. It is noted that the test set of compounds contained

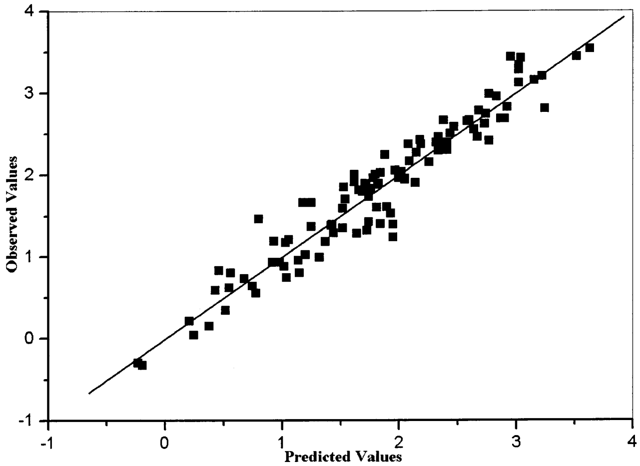

Figure 2. Relationship between logP calculated from eq 1 and experimental logP for the training set.

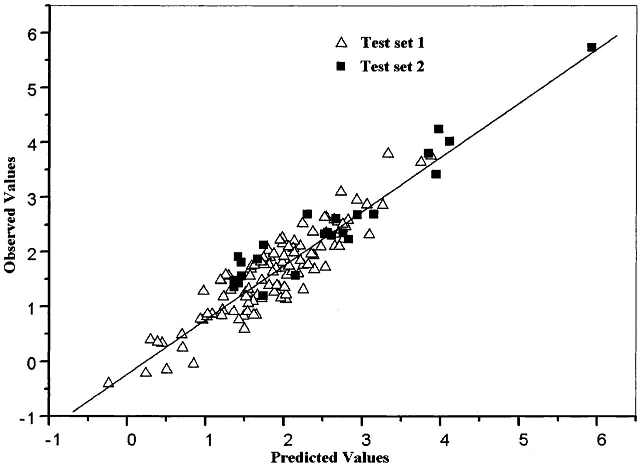

Figure 3. Relationship between logP calculated from eq 1 and experimental logP for the test set 1 (111 disubstituted benzenes) and the test set 2 (24 polysubstituted benzenes).

several substituted groups that were not presented in the training set, e.g., Br, F, $CONH_2$, and so on. Additionally, because the model was established based on the nonempirical molecular descriptors, it seems reasonable to believe that it also has a predictive power for the polysubstituted benzenes. The experimental and predicted logP values for these compounds as well as the residues are given in Table 2. The plot of predicted against experimental logP values is shown in Figure 3, where a good fit is observed.

As compared with meta- and para-disubsititued benzenes, ortho-disubsititued benzenes usually represent abnormal physicochemical properties as well as the partition between octanol and water. This can be ascribed to the interactions between two neighboring functional groups, especially the intramolecular hydrogen bonding. Equation 1 gives a good model for correlating and predicting the logP of disubstituted benzenes. However, it is still ambiguous how group−group interactions influence the logP of ortho-disubstituted benzenes, that is, which molecular descriptors in eq 1 are responsible for the group−group interactions. To address this problem, we treated the ortho-disubstituted benzenes separated from others in the training set and obtained the correlation for two subsets as follows.

meta and para

$$\log P = 0.12330\,\Sigma V_S^{-} + 0.33210\,V_{s,\max} + 0.04027\,V + 2.58200\,\mu^2/V - 0.02162\,\mathrm{PSA} + 0.09578$$

$$n = 68,\ \mathrm{sd} = 0.235,\ r = 0.967,\ \text{and}\ F = 181.43 \qquad (2)$$

ortho

$$\log P = 0.12527\,\Sigma V_S^{-} + 0.03619\,V + 1.91868\,\mu^2/V - 0.01315\,\mathrm{PSA} + 0.93709$$

$$n = 35,\ \mathrm{sd} = 0.296,\ r = 0.943,\ \text{and}\ F = 60.584 \qquad (3)$$

The correlation for ortho-disubstituted benzenes given in eq 3 is less satisfactory than that for meta- and para-disubstituted benzenes (eq 2) and that for the whole training set (eq 1). This is not beyond what one might expect and means that the influences of group−group interactions on partition are somewhat complicated. The $V_{s,\max}$ term, which is presented in eq 2, is found to be not statistically significant in the correlation for ortho-disubstituted benzenes. In addition, the $V_{\min}$ term, which gives minor but statistically significant contributions to logP, is not present in either eq 2 or eq 3. We speculate that these two molecular descriptors should be responsible for the interactions between two neighboring substituent groups. Interestingly, while correlating the differences between the computed quantities for the ortho-disubstituted benzenes and those for the meta- or para-disubstituted benzenes with their differences in logP, we have found a good three parameter relationship given in eq 4 (only the absolute value of ΔlogP over 0.4 log units, which is generally considered to be the experimental error range, is included). All of the three parameters ($\Delta V_{\min}$, $\Delta V_{s,\max}$, and ΔPSA) presented in this model are relevant to hydrogen bonding, which means that the variations in logP with the positions of substituent groups (i.e., the differences of logP between ortho- and meta- or para-disubstituted benzene isomers) basically resulted from their differences in the hydrogen bond-forming capacity. It is noted that the influences of group−group interactions on logP are represented through two different hydrogen-bonding styles. One is through intramolecular hydrogen bonding. In this case, the $V_{\min}$, $V_{s,\max}$, and PSA values for the ortho isomer all vary distinctly as compared with those for the meta and para isomers, e.g., for *m*-, *o*-, and *p*-hydroxylbenzaldehyde **52**−**54**. The other is through the variations of solvent accessible HBA and/or HBD (or intermolecular hydrogen bonding) where only the PSA value has distinct variation, e.g., for the isomers of nitroaniline **19**−**21**.

$$\Delta \log P = 0.96226\,\Delta V_{\min} + 0.23446\,\Delta V_{s,\max} - 0.16262\,\Delta\mathrm{PSA} - 0.33030$$

$$n = 17,\ \mathrm{sd} = 0.133,\ r = 0.980,\ \text{and}\ F = 106.16 \qquad (4)$$

In addition, we have found that only ΣV_S^{-} and V, which give the most important contributions in eqs 1−3, are statistically significant in the correlation for monosubstituted benzenes. Equation 5 gives the correlation for 19 monosubstituted benzenes. All of the substituent groups presented in this work are involved except benzylamine (an obvious outlier, and the linear correlation coefficient decreases to 0.912 when this solute is included). Accordingly, these two descriptors were thought to be mainly responsible for the variations in logP with substituted groups.

$$\log P = 0.17360\,\Sigma V_S^{-} + 0.02619\,V + 2.94625$$

$$n = 19,\ \mathrm{sd} = 0.291,\ r = 0.944,\ \text{and}\ F = 65.315 \qquad (5)$$

TABLE 2: Predictions of logP for Some Disubstituted and Polysubstituted Benzenes Using Eq 1

	logP (exp)	logP (calcd)	residual		logP (exp)	logP (calcd)	residual
			Disubstituted Benzenes				
3-bromo-1-nitrobenzene	2.64	2.54	0.10	3-ethylphenol	2.40	2.73	−0.33
4-bromo-1-nitrobenzene	2.55	2.69	−0.14	2-ethylphenol	2.47	2.79	−0.32
2-bromo-1-nitrobenzene	2.52	2.24	0.28	4-ethylphenol	2.58	2.82	−0.24
1,3-dibromobenzene	3.75	3.88	−0.13	3-methylbenzyl alcohol	1.60	2.19	−0.59
1,4-dibromobenzene	3.64	3.75	−0.11	4-methylbenzyl alcohol	1.58	2.04	−0.46
1,2-dibromobenzene	3.79	3.33	0.46	1,3-dimethoxybenzene	2.21	1.95	0.26
3-fluoro-1-nitrobenzene	1.90	1.82	0.08	1,4-dimethoxybenzene	2.03	1.80	0.23
4-fluoro-1-nitrobenzene	1.80	1.64	0.16	3-oxethylphenol	1.98	2.13	−0.15
2-fluoro-1-nitrobenzene	1.69	1.58	0.11	2-oxethylphenol	1.68	2.39	−0.71
1,2-difluorobenzene	2.37	2.37	0.00	4-oxethylphenol	1.81	1.98	−0.17
3-bromophenol	2.63	2.52	0.11	2-methylbenzaldehyde	2.09	2.06	0.03
4-bromophenol	2.59	2.64	−0.05	2-methoxybenzaldehyde	1.70	1.90	−0.20
2-bromophenol	2.35	2.54	−0.19	4-methoxybenzaldehyde	1.76	1.63	0.13
3-fluorophenol	1.93	1.77	0.16	3-cyanobenzaldehyde	1.18	1.23	−0.05
4-fluorophenol	1.77	1.81	−0.04	3-fluorophenoxyacetic acid	1.48	1.72	−0.24
2-fluorophenol	1.71	2.00	−0.29	2- fluorophenoxyacetic acid	1.39	1.81	−0.42
3-bromoaniline	2.10	2.15	−0.05	4- fluorophenoxyacetic acid	1.64	2.11	−0.47
2-bromoaniline	2.11	2.21	−0.10	3-trifluoromethylbenzyl alcohol	2.24	2.76	−0.52
4-bromoaniline	2.26	1.98	0.28	3-trifluoromethylacetophenone	2.10	2.71	−0.61
3-fluoroaniline	1.30	1.32	−0.02	3-carboxybenzaldehyde	1.76	2.25	−0.49
2-fluoroaniline	1.26	1.51	−0.25	2-methoxybenzyl alcohol	1.13	2.03	−0.90
4-fluoroaniline	1.15	1.72	−0.57	4-methoxybenzyl alcohol	1.10	1.61	−0.51
3-bromobenzoic acid	2.87	3.06	−0.19	3-hydroxybenzoic acid methyl ester	1.89	1.74	0.15
2-bromobenzoic acid	2.20	2.13	0.07	2-hydroxybenzoci acid methyl ester	1.96	1.87	0.09
4-bromobenzoic acid	2.86	3.26	−0.40	4-hydroxymethyl benzoic acid	0.93	1.22	−0.29
2-chlorobenzaldehyde	2.33	2.69	−0.36	3-fluoroacetophenone	1.77	2.31	−0.54
4-chlorobenzaldehyde	2.10	2.65	−0.55	4-fluoroacetophenone	1.72	2.53	−0.81
3-fluorobenzoic acid	2.15	1.98	0.17	4-chloroacetophenone	2.32	3.09	−0.77
2-fluorobenzoic acid	1.77	1.75	0.02	3-chloroacetophenone	2.51	2.75	−0.24
4-fluorobenzoic acid	2.07	2.12	−0.05	2-chloroacetophenone	2.09	2.47	−0.38
3-nitrobenzaldehyde	1.47	1.20	0.27	2-aminoacetophenone	1.63	1.86	−0.23
2-nitrobenzaldehyde	1.74	1.59	0.15	4-aminoacetophenone	0.83	1.49	−0.66
4-nitrobenzaldehyde	1.56	1.30	0.26	3-cyanoacetophenone	1.16	1.96	−0.80
3-chlorobenzyl alcohol	1.94	2.38	−0.44	4-cyanoacetophenone	1.22	2.03	−0.81
4-chlorobenzyl alcohol	1.96	2.35	−0.39	2-methoxyacetophenone	1.82	1.71	0.11
2-fluorobenzyl alcohol	1.31	2.25	−0.94	3-methoxyacetophenone	1.84	2.21	−0.37
4-fluorobenzyl alcohol	1.36	1.94	−0.58	4-methoxyacetophenone	1.74	2.07	−0.33
3-hydroxylbenzamide	0.39	0.30	0.09	4-nitrobenzylamine	1.06	1.55	−0.49
2-hydroxylbenzamide	1.28	0.98	0.30	3-fluorobenzamide	0.91	1.36	−0.45
4-hydroxylbenzamide	0.33	0.45	−0.12	2-fluorobenzamide	0.59	1.50	−0.91
2-nitrobenzyl alcohol	1.24	1.54	−0.30	4-fluorobenzamide	0.91	1.54	−0.63
3-nitrobenzyl alcohol	1.21	1.66	−0.45	2-nitrobenzamide	−0.15	0.51	−0.66
4-nitrobenzyl alcohol	1.26	1.88	−0.62	3-nitrobenzamide	0.77	0.94	−0.17
3-methoxyphenol	1.58	1.26	0.32	4-aminobenzamide	0.82	1.02	−0.20
4-methoxyphenol	1.32	1.55	−0.23	2-aminobenzamide	0.35	0.39	−0.04
3-hydroxybenzyl alcohol	0.49	0.70	−0.21	4-aminobenzamide	−0.41	−0.24	−0.17
4-hydroxybenzyl alcohol	0.25	0.71	−0.46	2-methylbenzamide	0.76	1.43	−0.67
3-aminobenzyl alcohol	−0.05	0.85	−0.90	3-methylbenzamide	1.18	1.52	−0.34
4-aminobenzyl alcohol	−0.22	0.24	−0.46	4-methylbenzamide	0.85	1.62	−0.77
3-trifluoromethylbenzoic acid	2.95	2.93	0.02	3-methoxybenzamide	0.84	1.21	−0.37
4-trifluoromethylbenzoic acid	3.10	2.73	0.37	2-methoxybenzamide	0.86	1.09	−0.23
3-cyanobenzoic acid	1.48	1.19	0.29	4-methoxybenzamide	0.86	1.03	−0.17
4-cyanobenzoic acid	1.56	1.57	−0.01	3-hydroxyacetophenone	1.39	1.91	−0.52
3-hydroxyphenylacetic acid	0.85	1.54	−0.69	2-hydroxyacetophenone	1.92	2.00	−0.08
2-hydroxyphenylacetic acid	0.85	1.65	−0.80	4-hydroxyacetophenone	1.35	2.01	−0.66
4-hydroxyphenylacetic acid	0.75	0.97	−0.22				
			Polysubstituted Benzenes				
2,4-dichlorotoluene	4.24	3.98	0.26	4-methoxy-2-nitroaniline	1.86	1.67	0.19
1,2,4-trichlorobenzene	4.02	4.12	−0.10	2,5-dimethoxyaniline	1.20	1.74	−0.54
hexachlorobenzene	5.73	5.93	−0.20	2-chloro-6-nitrobenzoic acid	1.56	1.47	0.09
3,4-dichloroaniline	2.69	2.30	0.39	2-nitro-5-chlorobenzoic acid	2.13	1.75	0.38
2,3-xylenol	2.61	2.67	−0.06	2-amino-5-chlorobenzoic acid	1.57	2.15	−0.58
2,4-xylenol	2.30	2.61	−0.31	3,5-dichloroanisole	3.80	3.85	−0.05
2,5-xylenol	2.33	2.52	−0.19	2-amino-4-nitrobenzoic acid	1.91	1.42	0.49
2,6-xylenol	2.36	2.56	−0.20	4-amino-2-methoxybenzoic acid	1.35	1.37	−0.02
3,4-xylenol	2.23	2.83	−0.60	2-methyl-3-nitroanisole	2.68	2.94	−0.26
3,5-xylenol	2.35	2.76	−0.41	4-hydroxy-3-methoxybenzoic acid	1.43	1.43	0.00
2-amino-4-chlorophenol	1.81	1.46	0.35	2-chloro-4-hydroxy-3-methoxybenzene	2.69	3.15	−0.46
2-methoxy-5-nitroaniline	1.47	1.37	0.10	1,3,5-trimethylbenzene	3.42	3.95	−0.53

Theoretical linear solvation energy relationship (TLSER) descriptors are thought to be excellent theoretical descriptors for reflecting solute−solvent interactions[27,44−46] and have been used successfully to correlate and predict the partition between octanol and water.[27] However, a poor correlation with a correlation coefficient of only 0.69 was observed when they were applied in our present data set. Considering that the TLSER descriptors and those used in present work are not of equal quality, i.e., the former is calculated at semiempirical (MNDO) SCF level, whereas the latter is calculated at the HF/6-31G* level, we calculated these descriptors at a higher level and obtained an improved but far from satisfactory correlation ($r = 0.765$). Although Politzer et al.[47] found that $V_{s,max}$ and $V_{s,min}$ correlated well to the q^+ (the most positive charge on a hydrogen atom) and q^- (the most negative charge), respectively, substitution of q^+ and q^- for $V_{s,max}$ and V_{min} would make the correlation worse. This implies that atomic partial charges are not as sensitive to group−group interactions (also intramolecular HB) as $V_{s,max}$ and V_{min}, in some cases, and they even give wrong results. For example, the q^+ values for *m*-, *o*-, and *p*-hydroxylbenzaldehyde (**52**−**54**) calculated at HF/6-31G* are 0.46206, 0.51463, and 0.46238, respectively, and q^- values are −0.75545, −0.76625, and −0.74485. The HBA basicity and HBD acidity of *o*-hydroxylbenzaldehyde that can form an intramolecular hydrogen bond, as reflected in the q^+ and q^-, do not lower as one might expect but rather increase. It is also the reason that only middle correlations ($r = 0.750$ for the AM1 level and $r = 0.849$ for the HF/6-31G* level) were found when we applied Bodor's theoretical descriptors[48,49] in our present data set. It is necessary to point out that there are two size-dependent descriptors Σ_S^- and PSA introduced in eq 1 apart from the cavity term V. This is quite sound and consistent with the fact that logP has additive character. Perhaps it is another reason that TLSER descriptors fail to be applied in more complicated compounds. A more detailed test of other computed descriptors is not the object of this work.

4. Conclusion

A good correlation between logP and computed descriptors for a training set of 103 disubstituted benzenes given by eq 1 has been established and shows strong predictive ability. Of these descriptors introduced in eq 1, ΣV_S^-, a new descriptor proposed in this work and defined as the sum of the surface minima values of the electrostatic potential, gives the most significant contribution to the logP. This descriptor, together with the molecular volume V, shown in the correlation for monosubstituted benzenes given by eq 5, is mainly responsible for variations in logP with the substituent groups. The influence of group−group interactions on the logP values is basically thought to be relevant to hydrogen bonding and represented mainly by variations of the V_{min}, $V_{s,max}$, and PSA values of isomers.

The TLSER approach, and the one proposed by Bodor, have been demonstrated to not be suited for correlating or predicting the logP for disubstituted benzenes. It may be because the partial atomic charges are not as sensitive to those group−group interactions as the descriptors derived from electrostatic potential and/or size-dependent descriptors are unavailable in these approaches. Perhaps other molecular descriptors, such as those calculated from the CODESSA (comprehensive descriptors for statistical and structural analysis) program,[50,51] may result in better correlation relationships, but it may also well be that these descriptors cannot give as explicit physical meaning as those ones used in the present work.

Acknowledgment. We thank the National Natural Science Foundation of China (No. 20173050) for financial support.

References and Notes

(1) Sangster, J. *Octanol−Water Partition Coefficients, Fundamentals and Physical Chemistry*; Wiley Series in Solution Chemistry, Vol. 2; Wiley: New York, 1997.

(2) Fujita, T.; Iwasa, J.; Hansch, C. *J. Am. Chem. Soc.* **1964**, *86*, 5175.

(3) Leo, A. *J. Chem. Rev.* **1993**, *93*, 1281.

(4) Buchwald, P.; Bodor, N. *Curr. Med. Chem.* **1998**, *5*, 353.

(5) Katritzky, A. R.; Maran, U.; Lobanov, V. S.; Karelson, M. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1.

(6) Mannhold, R.; Waterbeemd, H. V. D. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 337.

(7) Fujita, T. *Prog. Phys. Org. Chem.* **1983**, *14*, 75.

(8) Brandstrom, A. *Acta Pharm. Suec.* **1982**, *19*, 175.

(9) Leo, A. *J. Chem. Soc., Perkin Trans. 2* **1983**, 825.

(10) Brandstrom, A. *J. Chem. Soc., Perkin Trans. 2* **1999**, 2419.

(11) Camilleri, P.; Watts, S. A.; Boraston, J. A. *J. Chem. Soc., Perkin Trans. 2* **1988**, 1699.

(12) Haeberlein, M.; Brinck, T. *J. Chem. Soc., Perkin Trans. 2* **1997**, *2*, 289.

(13) Breindl, A.; Beck, B.; Clark, T.; Glen, R. C. *J. Mol. Model.* **1997**, *3*, 142.

(14) Brinck, T.; Murray, J. S.; Politzer, P. *J. Org. Chem.* **1993**, *58*, 7070.

(15) Murray, J. S.; Abu-Awwad, F.; Politzer, P. *J. Phys. Chem. A* **1999**, *103*, 1853.

(16) Hagelin, H.; Murray, J. S.; Brinck, T.; Berthelot, M.; Politzer, P. *Can. J. Chem.* **1995**, *73*, 483.

(17) Murray, J. S.; Brinck, T.; Lane, P.; Paulsen, K.; Politzer, P. *J. Mol. Struct. (THEOCHEM)* **1994**, *307*, 55.

(18) Zou, J. W.; Yu, Q. S.; Shang, Z. C. *J. Chem. Soc., Perkin Trans. 2* **2001**, *8*, 1439.

(19) Politzer, P.; Murray, J. S.; Peralta-Inga, Z. *Int. J. Quantum Chem.* **2001**, *85*, 676.

(20) Murray, J. S.; Politzer, P. *J. Mol. Struct. (THEOCHEM)* **1998**, *425*, 107.

(21) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseris, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. GAUSSIAN 98, Revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.

(22) Hehre, W. J.; Radom, R.; Schleyre, P. V. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley: New York, 1986.

(23) Shang, Z. C.; Yu, Q. S.; Lin, R. S. *Acta Phys. Chim. Sinica* **1997**, *13*, 1097.

(24) Hansch, C.; Leo, A. *Exploring QSAR: Hydropholic, Electronic, and Steric Constants*; ACS: Washington, DC, 1995.

(25) Carl, L. Y.; Sachin, N.; Li, B.; Deepa, R. B.; Saumya, T. Solubility in Water and Octanol−Water Partition Coefficient. In *Chemical Properties Handbook*; Carl, L. Y., Ed.; McGraw-Hill: Singapore, 1999, p 364.

(26) EI Tayar, N.; Testa, B.; Carrupt, P. A. *J. Phys. Chem.* **1992**, *96*, 1455.

(27) Famini, G. R.; Penski, C. A.; Wilson, L. Y. *J. Phys. Org. Chem.* **1992**, *5*, 395.

(28) Kamlet, M. J.; Doherty, R. M.; Carr, P.; Abraham, M. H.; Marcus, Y.; Taft, R. W. *J. Phys. Chem.* **1988**, *92*, 5244.

(29) Bodor, N.; Buchwald, P. *J. Phys. Chem. B* **1997**, *101*, 3404.

(30) Leahy, D. E.; Morris, J. J.; Taylor, P. J.; Wait, A. R. *J. Chem. Soc., Perkin Trans. 2* **1992**, 705.

(31) Brinck, T.; Murray, J. S.; Politzer, P. *Mol. Phys.* **1992**, *76*, 609.

(32) Van de Waterbeemd, H.; Camenisch, G.; Folkers, G.; Raevsky, O. A. *Quant. Struct.−Act. Relat.* **1996**, *15*, 480.

(33) Clark, D. E. *J. Pharm. Sci.* **1999**, *88*, 807.

(34) Clark, D. E. *J. Pharm. Sci.* **1999**, *88*, 815.

(35) Palm, K.; Luthman, K.; Ungell, A. L.; Strandlund, G.; Artursson, P. *J. Pharm. Sci.* **1996**, *85*, 32.

(36) Krarup, L. H.; Christensen, I. T.; Hovgaard, L.; Frokjaer S. *Pharm. Res.* **1998**, *15*, 972.

(37) Van de Waterbeemd, H.; Camenisch, G.; Folkers, G.; Chretien, J. R.; Raevsky, O. A. *J. Drug Targeting* **1998**, *6*, 151.

(38) Palm, K.; Artursson, P.; Luthman, K. Experimental and theoretical predictions of intestinal drug absorption. In *Computer-Assisted Lead Finding*

and Optimization; Van de Waterbeemd, H., Testa, B., Folkers, G., Eds.; Wiley-VCH: Weinheim, 1997; p 227.

(39) Osterberg, T.; Norinder, U. *J. Chem. Inf. Comput Sci.* **2000**, *40*, 1408.

(40) Hansch, C.; Leo, A. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*; ACS: Washington, DC, 1995.

(41) Simkin, B. Y.; Sheikhet, I. I. *Quantum Chemical and Statistical Theory of Solutions: A Computational Approach*; Ellis Horwood: London, 1995.

(42) Mu, L.; Drago, R. S.; Richardson, D. E. *J. Chem. Soc., Perkin Trans. 2* **1998**, 159.

(43) Belesley, D. A.; Kuh, E.; Welsh, R. E. *Regression Diagnostics*; Wiley: New York, 1990.

(44) Wilson, L. Y.; Famini, G. R. *J. Med. Chem.* **1991**, *34*, 1668.

(45) Headley, A. D.; Starnes, S. D.; Wilson, L. Y.; Famini, G. R. *J. Org. Chem.* **1994**, *59*, 8040.

(46) Famini, G. R.; Wilson, L. Y. *J. Phys. Org. Chem.* **1999**, *12*, 645.

(47) Murray, J. S.; Politzer, P.; Famini, G. R. *J. Mol. Struct. (THEOCHEM)* **1998**, *454*, 299.

(48) Bodor, N.; Gabanyi, Z.; Wong, C. K. *J. Am. Chem. Soc.* **1989**, *111*, 3783.

(49) Bodor, N.; Huang, M. J. *J. Pharm. Sci.* **1992**, *81*, 272.

(50) Katrizky, A. R.; Tamm, T.; Wang, Y.; Karelson, M. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 692.

(51) Katrizky, A. R.; Tamm, T.; Wang, Y.; Sild, S.; Karelson, M. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 684.