

Molecular Polarizability. 1. Relationship to Water Solubility of Alkanes and Alcohols

Cao Chenzhong^{†,‡,§} and Li Zhiliang^{*,†,||}

Department of Chemistry and Chemical Engineering, Institute of Chemometrics and Pharmacy (ICP), Hunan University, Changsha 410082, P.R. China, Department of Chemistry, Xiangtan Teacher's College, Xiangtan 411100, P.R. China, and College of Chemistry and Chemical Engineering, Institute of Molecular Pharmacy (IMP), Chongqing University, Chongqing 630044, P.R. China

Received December 18, 1996[®]

The polarizability effect index (PEI) of groups and the molecular polarizability effect index (MPEI) are proposed on the basis of the principle that molecules are polarized by electric fields. Furthermore, by taking the sum of bond lengths (SBL) in a molecule and the difference in molecular polarizability effect index values (Δ MPEI) between branching chain and normal chain isomers as parameters, we studied the correlations of cavity surface area (CSA), water solubility ($-\log S$), and partition coefficients in *n*-octanol/water ($\log P$) with the SBL and Δ MPEI parameters for the alkanes and alcohols and got excellent linear correlations.

1. INTRODUCTION

Quantitative structure–activity (structure–property) relationships (QSAR/QSPR) is an interesting field. Solving this problem even for a narrow class of compounds is of great importance, since the regularities disclosed could be available for use in forecasting the physical or chemical properties and predicting biological activities of other compounds belonging to a given class, in developing a new theory, and in explaining the phenomena observed for the compounds. To find quantitative relationships between structures and properties, various methods of mathematical modeling have been applied.^{1–3} In most of the mathematical modeling, the topological indices were taken as molecular structure parameters; there are more than 100 kinds of topological indices developed to date.² The Wiener index,⁴ Randić index,⁵ Balaban index⁶ and Kier index⁷ are all widely applied. However, it is difficult to provide physical insight into the molecular behavior for a purely topological representation of structure in many cases. Recently, more attention has been paid to the development of physical descriptors such as electronic descriptors⁸ and geometrical descriptors.⁹ These descriptors could aid the understanding of molecular properties. Hall¹⁰ et al. combined electrotopological state indices and other indices and obtained a good correlation between these indices and the boiling points of alkanes and alcohols.

As is well-known, the molecular geometrical shape and electronic potential are feature properties related to molecular structure. In addition, the molecular polarizability is also a feature property related to molecular structure. Therefore, it can be predicted that descriptors derived from the molecular polarizability might be applied widely in QSPR/QSAR research.

Table 1. PEI and Δ PEI Values of Normal Alkyl $H(CH_2)_N$

<i>N</i>	PEI	<i>N</i>	PEI	<i>N_i</i>	Δ PEI	<i>N_i</i>	Δ PEI
1	1.0000	11	1.2551	1	1.000 000	11	0.002 375
2	1.1405	12	1.2571	2	0.140 526	12	0.001 972
3	1.1887	13	1.2587	3	0.048 132	13	0.001 628
4	1.2122	14	1.2602	4	0.025 503	14	0.001 421
5	1.2260	15	1.2614	5	0.013 800	15	0.001 229
6	1.2350	16	1.2625	6	0.009 052	16	0.001 073
7	1.2414	17	1.2634	7	0.006 388	17	0.000 945
8	1.2461	18	1.2642	8	0.004 748	18	0.000 838
9	1.2498	19	1.2650	9	0.003 666	19	0.000 749
10	1.2527	20	1.2657	10	0.002 916	20	0.000 673

In this work, a descriptor called the molecular polarizability effect index (MPEI) is proposed on the basis of the principle of a molecule being polarized in an electric field, and it is applied in QSPR/QSAR research.

2. MOLECULAR POLARIZABILITY EFFECT INDEX

As is well-known, the stabilizing energy caused by the polarizability effect for a substituent R interacting with a point charge q is¹¹

$$E_{(R)} = -\alpha q^2 / (2Dr^4) \quad (1)$$

where α is the polarizability of the substituent R, D is the effective dielectric constant, and r is the distance of the point charge q to the center of substituent R. The key is to obtain the values of parameters α and r for the application of eq 1.

In this paper, taking the alkyl substituent as an example, we discuss the calculation of stabilizing energy (eq 1). On the basis of the method for additivity of molecular polarizability,¹⁰ the total stabilizing energy $E_{(R)}$ (eq 1) of an alkyl substituent R can be taken as the sum of contributions from essential units. The essential units are four kinds of CH_3 , CH_2- , $CH<$, and $-C<$, which consist of three C–H bonds, two C–H and one C–C bonds, one C–H and two C–C bonds, and three C–C bonds, respectively. Let α_i be the polarizability of i th essential unit and r_i the average distance

[†] Hunan University.

[‡] Xiangtan Teacher's College.

[§] Visiting Professor on leave from Xiangtan Teacher's College.

^{||} Chongqing University.

* Corresponding author.

[®] Abstract published in *Advance ACS Abstracts*, December 15, 1997.

Table 2. MPEI Value of Alkanes Bearing 1–10 Carbon Atoms

no.	alkane	MPEI	no.	alkane	MPEI	no.	alkane	MPEI
1	methane	1	61	2,3-dimethylheptane	12.7268	121	2,3,5-trimethylheptane	14.3858
2	ethane	2.2811	62	3-ethyl-2-methylhexane	12.7854	122	2,5-dimethyl-3-ethylhexane	14.4154
3	propane	3.6584	63	3,4-dimethylheptane	12.7705	123	2,4,5-trimethylheptane	14.3953
4	2-methylpropane	5.1320	64	3-ethyl-4-methylhexane	12.8143	124	2,3,6-trimethylheptane	14.3367
5	butane	5.0827	65	2,4-dimethylheptane	12.7167	125	2,3-dimethyloctane	14.2414
6	2,2-dimethylpropane	6.7018	66	4-ethyl-2-methylhexane	12.7565	126	3-ethyl-2-methylheptane	14.3118
7	2-methylbutane	6.6033	67	3,5-dimethylheptane	12.7416	127	3,4-dimethyloctane	14.2937
8	pentane	6.5346	68	2,5-dimethylheptane	12.6979	128	4-isopropylheptane	14.3320
9	2,2-dimethylbutane	8.2201	69	2,6-dimethylheptane	12.6584	129	4-ethyl-3-methylheptane	14.3608
10	2,3-dimethylbutane	8.1709	70	2-methyloctane	12.5683	130	4,5-dimethyloctane	14.3085
11	2-methylpentane	8.0828	71	3-methyloctane	12.6058	131	3-ethyl-4-methylheptane	14.3555
12	3-methylpentane	8.1022	72	4-methyloctane	12.6063	132	3,4-diethylhexane	14.4046
13	hexane	8.0046	73	3-ethylheptane	12.6581	133	2,4,6-trimethylheptane	14.3361
14	2,2,3-trimethylbutane	9.8347	74	4-ethylheptane	12.6730	134	2,4-dimethyloctane	14.2366
15	2,2-dimethylpentane	9.7273	75	nonane	12.4794	135	4-ethyl-2-methylheptane	14.2975
16	3,3-dimethylpentane	9.7661	76	2,2,3,3,4-pentamethylpentane	14.8442	136	3,5-dimethyloctane	14.2741
17	2,3-dimethylpentane	9.6794	77	2,2,3,3-tetramethylhexane	14.7081	137	3-ethyl-5-methylheptane	14.3171
18	2,4-dimethylpentane	9.6586	78	3-ethyl-2,2,3-trimethylpentane	14.7755	138	2,5-dimethyloctane	14.2273
19	2-methylhexane	9.5709	79	3,3,4,4-tetramethylhexane	14.7659	139	5-ethyl-2-methylheptane	14.2681
20	3-methylhexane	9.5998	80	2,2,3,4,4-pentamethylpentane	14.7858	140	3,6-dimethyloctane	14.2500
21	3-ethylpentane	9.6287	81	2,2,3,4-tetramethylhexane	14.6201	141	2,6-dimethyloctane	14.2071
22	heptane	9.4874	82	3-ethyl-2,2,4-trimethylpentane	14.6486	142	2,7-dimethyloctane	14.1663
23	2,2,3,3-tetramethylbutane	11.5456	83	2,3,4,4-tetramethylhexane	14.6493	143	2-methylnonane	14.0753
24	2,2,3-trimethylpentane	11.3889	84	2,2,3,5-tetramethylhexane	14.5527	144	3-methylnonane	14.1149
25	2,3,3-trimethylpentane	11.4082	85	2,2,3-trimethylheptane	14.4491	145	4-methylnonane	14.1330
26	2,2,4-trimethylpentane	11.3306	86	2,2-dimethyl-3-ethylhexane	14.5225	146	3-ethyloctane	14.1727
27	2,2-dimethylhexane	11.2335	87	3,3,4-trimethylheptane	14.5271	147	5-methylnonane	14.1384
28	3,3-dimethylhexane	11.2913	88	3,3-dimethyl-4-ethylhexane	14.5803	148	4-ethyloctane	14.1961
29	3-ethyl-3-methylpentane	11.3396	89	2,3,3,4-tetramethylhexane	14.6684	149	4-propylheptane	14.2143
30	2,3,4-trimethylpentane	11.3202	90	3,4,4-trimethylheptane	14.5366	150	decane	13.9848
31	2,3-dimethylhexane	11.2131	91	3,4-dimethyl-3-ethylhexane	14.5997			
32	3-ethyl-2-methylpentane	11.2515	92	3-ethyl-2,3,4-trimethylpentane	14.7069			
33	3,4-dimethylhexane	11.2420	93	2,3,3,5-tetramethylhexane	14.5911			
34	2,4-dimethylhexane	11.1937	94	2,3,3-trimethylpentane	14.4834			
35	2,5-dimethylhexane	11.1553	95	2,3-dimethyl-3-ethylhexane	14.5708			
36	2-methylheptane	11.0665	96	3,3-diethyl-2-methylpentane	14.6381			
37	3-methylheptane	11.1007	97	2,2,4,4-tetramethylhexane	14.6305			
38	4-methylheptane	11.1102	98	2,2,4,5-tetramethylhexane	14.5333			
39	3-ethylhexane	11.1445	99	2,2,4-trimethylheptane	14.4198			
40	octane	10.9797	100	2,2-dimethyl-4-ethylhexane	14.4647			
41	2,2,3,3-tetramethylpentane	13.1467	101	3,3,5-trimethylheptane	14.4788			
42	2,2,3,4-tetramethylpentane	13.0392	102	2,4,4-trimethylheptane	14.4633			
43	2,2,3-trimethylhexane	12.9227	103	2,4-dimethyl-4-ethylhexane	14.5244			
44	2,2-dimethyl-3-ethylpentane	12.9706	104	2,2,5,5-tetramethylhexane	14.5346			
45	3,3,4-trimethylhexane	12.9805	105	2,2,5-trimethylheptane	14.3913			
46	2,3,3,4-tetramethylpentane	13.0781	106	2,5,5-trimethylheptane	14.4256			
47	2,3,3-trimethylhexane	12.9516	107	2,2,6-trimethylheptane	14.3463			
48	2,3-dimethyl-3-ethylpentane	13.0094	108	2,2-dimethyloctane	14.2531			
49	2,2,4,4-tetramethylpentane	13.0303	109	3,3-dimethyloctane	14.3281			
50	2,2,4-trimethylhexane	12.8839	110	4,4-dimethyloctane	14.3578			
51	2,4,4-trimethylhexane	12.9128	111	3-ethyl-3-methylheptane	14.4147			
52	2,2,5-trimethylhexane	12.8360	112	4-ethyl-4-methylheptane	14.4390			
53	2,2-dimethylheptane	12.7418	113	3,3-diethylhexane	14.5021			
54	3,3-dimethylheptane	12.8103	114	2,3,4,5-tetramethylhexane	14.5418			
55	4,4-dimethylheptane	12.8293	115	2,3,4-trimethylheptane	14.4242			
56	3-ethyl-3-methylhexane	12.8829	116	2,3-dimethyl-4-ethylhexane	14.4732			
57	3,3-diethylpentane	12.9407	117	2,4-dimethyl-3-ethylhexane	14.4827			
58	2,3,4-trimethylhexane	12.8829	118	3,4,5-trimethylheptane	14.4584			
59	2,4-dimethyl-3-ethylpentane	12.9019	119	2,4-dimethyl-3-isopropylpentane	14.5799			
60	2,3,5-trimethylhexane	12.8251	120	3-isopropyl-2-methylhexane	14.4538			

from the point charge (q) to the i th carbon atom. Thus eq 1 can be expressed as

$$E_{(R)} = [-q^2/(2D)][\sum(\alpha_i/r_i^4)] \quad (2)$$

That is, the total stabilizing energy $E_{(R)}$ of alkyl substituent R is the sum of the stabilizing energies resulting from its essential unit. In eq 2, the polarizability value α_i of the i th essential unit can be obtained by taking the sum of the polarizabilities of all chemical bonds in this unit. That is,

CH_3 , CH_2- , $\text{CH}<$, and $-\text{C}<$, the α_i values are $\alpha(\text{CH}_3) = 3\alpha(\text{C}-\text{H})$, $\alpha(\text{CH}_2-) = 2\alpha(\text{C}-\text{H}) + \alpha(\text{C}-\text{C})$, $\alpha(\text{CH}<) = \alpha(\text{C}-\text{H}) + 2\alpha(\text{C}-\text{C})$ and $(-\text{C}<) = 3\alpha(\text{C}-\text{C})$, respectively. According to Miller's¹² work, we obtain $\alpha(\text{C}-\text{H}) = 1/4\alpha_{\text{CTE}} + \alpha_{\text{H}}$, $\alpha(\text{C}-\text{C}) = 2(1/4\alpha_{\text{CTE}})$. From the experimental values¹² of H atoms, $\alpha_{\text{H}} = 0.358 \text{ \AA}^3$ and C atoms, with sp^3 hybridization, $\alpha_{\text{CTE}} = 1.443$, we obtain $\alpha(\text{C}-\text{H}) = 0.719$ and $\alpha(\text{C}-\text{C}) = 0.721 \text{ \AA}^3$, that is $\alpha(\text{C}-\text{H}) \approx \alpha(\text{C}-\text{C}) = 0.72 \text{ \AA}^3$. Therefore, the polarizability values of all four essential

Table 3. Water Solubilities (log *S*), Cavity Surface Area (CSA), Sum of Bond Length (SBL), and Difference of Molecular Polarizability Index (Δ MPEI) for Alcohol and Hydrocarbons

no.	compd	SBL	Δ MPEI	$-\log S_{\text{obs}}^a$	$-\log S_{\text{calc}}^b$	CSA _{lit} ¹⁵	CSA _{calc}
1	1-butanol	16.7400	0.0000	-0.00	-0.13	272.10	268.56
2	2-methyl-1-propanol	16.7400	0.0687	-0.01	-0.24	263.80	261.73
3	2-butanol	16.7400	0.0687	-0.03	-0.24	264.10	261.73
4	1-pentanol	20.4400	0.0000	0.58	0.47	303.90	300.59
5	3-methyl-1-butanol	20.4400	0.0782	0.51	0.35	291.40	292.82
6	2-methyl-1-butanol	20.4400	0.0976	0.46	0.32	289.40	290.89
7	2-pentanol	20.4400	0.0782	0.28	0.35	295.90	292.82
8	3-pentanol	20.4400	0.0976	0.21	0.32	293.50	290.89
9	3-methyl-2-butanol	20.4400	0.1662	0.18	0.22	284.30	284.07
10	2-methyl-2-butanol	20.4400	0.2155	-0.15	0.14	282.50	279.17
11	1-hexanol	24.1400	0.0000	1.21	1.08	335.70	332.63
12	2-hexanol	24.1400	0.0835	0.87	0.95	327.70	324.33
13	3-hexanol	24.1400	0.1124	0.80	0.91	325.30	321.45
14	3-methyl-3-pentanol	24.1400	0.2786	0.36	0.65	305.80	304.93
15	2-methyl-2-pentanol	24.1400	0.2398	0.49	0.71	314.30	308.79
16	2-methyl-3-pentanol	24.1400	0.2100	0.70	0.76	314.30	311.75
17	3-methyl-2-pentanol	24.1400	0.2100	0.71	0.76	311.30	311.75
18	2,3-dimethyl-2-butanol	24.1400	0.3473	0.37	0.55	301.20	298.10
19	3,3-dimethyl-1-butanol	24.1400	0.2398	1.12	0.71	307.50	308.79
20	3,3-dimethyl-2-butanol	24.1400	0.3473	0.61	0.55	296.70	298.10
21	4-methyl-1-pentanol	24.1400	0.0835	0.99	0.95	323.00	324.33
22	4-methyl-2-pentanol	24.1400	0.1711	0.79	0.82	314.90	315.62
23	2-ethyl-1-butanol	24.1400	0.1413	1.21	0.86	308.60	318.58
24	cyclohexanol	23.5200	0.5127	0.42	0.19	290.50	276.28
25	1-heptanol	27.8400	0.0000	1.81	1.68	367.50	364.66
26	2-methyl-2-hexanol	27.8400	0.2537	1.07	1.29	346.10	339.44
27	3-methyl-3-hexanol	27.8400	0.3116	0.98	1.21	337.70	333.68
28	3-ethyl-3-pentanol	27.8400	0.3599	0.83	1.13	324.40	328.88
29	2,3-dimethyl-2-pentanol	27.8400	0.4091	0.87	1.06	323.80	323.99
30	2,3-dimethyl-3-pentanol	27.8400	0.4285	0.84	1.03	321.80	322.06
31	2,4-dimethyl-2-pentanol	27.8400	0.3509	0.93	1.14	328.60	329.77
32	2,4-dimethyl-3-pentanol	27.8400	0.3405	1.22	1.16	331.70	330.81
33	2,2-dimethyl-3-pentanol	27.8400	0.4091	1.15	1.06	326.10	323.99
34	3-heptanol	27.8400	0.1210	1.39	1.50	357.10	352.63
35	4-heptanol	27.8400	0.1305	1.39	1.48	357.10	351.69
36	1-octanol	31.5400	0.0000	2.35	2.29	399.40	396.70
37	2,2,3-trimethyl-3-pentanol	31.5400	0.6674	1.27	1.26	335.20	330.34
38	2-octanol	31.5400	0.0889	2.06	2.15	391.00	387.86
39	2-ethyl-1-hexanol	31.5400	0.1788	2.17	2.01	371.30	378.92
40	1-nonanol	35.2400	0.0000	3.00	2.89	431.20	428.73
41	2-nonanol	35.2400	0.0904	2.74	2.75	423.20	419.74
42	3-nonanol	35.2400	0.1301	2.66	2.69	420.80	415.79
43	4-nonanol	35.2400	0.1482	2.59	2.67	420.80	414.00
44	5-nonanol	35.2400	0.1535	2.49	2.66	420.80	413.47
45	2,6-dimethyl-4-heptanol	35.2400	0.2680	2.51	2.48	394.00	402.08
46	3,5-dimethyl-4-heptanol	35.2400	0.4736	2.30	2.17	379.30	381.64
47	2,2-diethyl-1-pentanol	35.2400	0.5173	2.42	2.10	372.50	377.30
48	7-methyl-1-octanol	35.2400	0.0904	2.49	2.75	418.70	419.74
49	3,5,5-trimethyl-1-hexanol	35.2400	0.4350	2.51	2.23	376.60	385.48
50	1-decanol	38.9400	0.0000	3.70	3.50	463.00	460.76
51	1-dodecanol	46.3400	0.0000	4.64	4.71	527.00	524.83
52	<i>n</i> -butane	15.4200	0.0000	2.63	2.73	255.20	257.13
53	2-methylpropane	15.4200	0.0493	2.55	2.66	249.10	252.23
54	<i>n</i> -pentane	19.1200	0.0000	3.27	3.34	287.00	289.16
55	2-methylbutane	19.1200	0.0687	3.18	3.23	274.60	282.33
56	3-methylbutane	22.8200	0.0976	3.83	3.79	300.10	311.50
57	neopentane	19.1200	0.1672	3.13	3.08	270.10	272.54
58	2,2-dimethylbutane	22.8200	0.2155	3.67	3.61	290.80	299.77
59	2,4-dimethylpentane	26.5200	0.1711	4.39	4.28	324.70	336.22
60	2,2,4-trimethylpentane	30.2200	0.3509	4.67	4.61	338.90	350.38
61	2,2,5-trimethylhexane	33.9200	0.3566	5.05	5.21	373.00	381.85
62	cyclohexane	22.2000	0.4003	3.18	3.23	279.10	276.03
63	methylcyclohexane	25.9000	0.5127	3.85	3.66	304.90	296.89
64	1- <i>cis</i> -2-dimethylcyclohexane	29.6000	0.6627	4.27	4.03	315.50	314.01
65	cycloheptane	25.9000	0.4828	3.52	3.70	301.90	299.86
66	cyclooctane	29.6000	0.5253	4.15	4.24	322.60	327.67
67	<i>n</i> -hexane	22.8200	0.0000	3.95	3.94	319.00	321.20
68	<i>n</i> -heptane	26.5200	0.0000	4.53	4.55	351.00	353.23
69	<i>n</i> -octane	30.2200	0.0000	5.24	5.15	383.00	385.27

^a From Ref 16. ^b Based on eq 9. ^c D_{OH} is taken as 1 for alcohols and as 0 for alkanes.

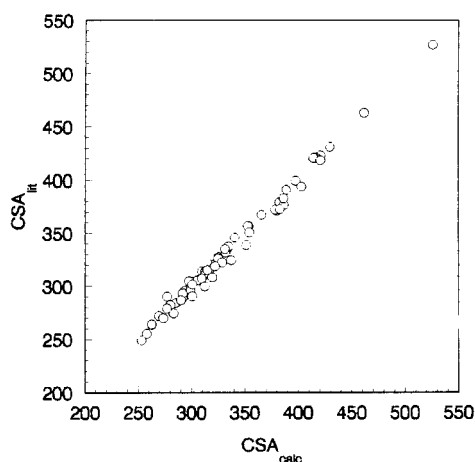


Figure 1. Plot of CSA_{lit} vs CSA_{calc} for alcohols and hydrocarbons.

Table 4. Predicted CSA Values of Some Alkanes

compd	SBL	Δ MEI	CSA_{lit}^a	CSA_{calc}^b
methane	4.32	0	152.36	161.03
ethane	8.02	0	191.52	193.06
propane	11.72	0	223.35	225.10
2-methylpentane	22.82	0.0782	306.40	313.43
cyclopentane	18.50	0.3520	255.40	248.80
methylcyclopentane	22.20	0.4496	282.71	271.13

^a From ref 14. ^b Based on eq 9.

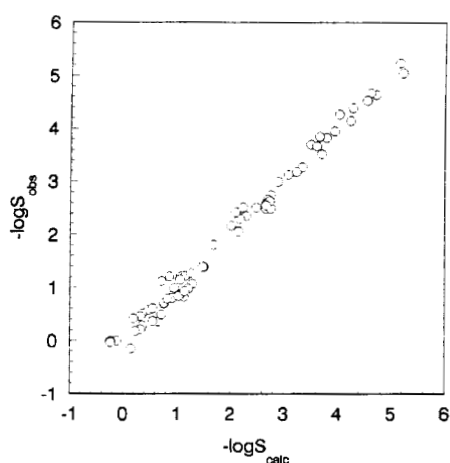


Figure 2. Plot of $-\log S_{obs}$ vs $-\log S_{calc}$ for alcohols and hydrocarbons.

units CH_3 , CH_2- , $CH<$, and $-C<$ are approximately equal to 2.16 \AA^3 . Here we take α_i as a constant, $\alpha_i = 2.16 \text{ \AA}^3$, to simplify the calculation.

The r_i in eq 2 can be computed as follows. In the carbon chain of an alkyl substituent, if the rejection among its essential units is ignored and the carbon chain is assumed to have freedom of rotation, the expression of average distance r_i between the point charge (q) and i th carbon atom can be obtained on the basis of the freedom rotation chain model¹³

$$r_i^2 = l^2 \left[N_i \frac{1 + \cos \theta}{1 - \cos \theta} - \frac{2 \cos \theta (1 - \cos^{N_i} \theta)}{(1 - \cos \theta)^2} \right] \quad (3)$$

where l is the length of C–C bond, N_i is the carbon atom number from the point charge (q) to the i th essential unit, and θ is the supplementary angle of bond angle $\angle CCC$ (that is, $\theta = 180^\circ - 109.5^\circ = 70.5^\circ$ for sp^3 hybridization).

Table 5. Observed and Predicted Solubilities ($\log S$) for Some Paraffins

no.	compd	SBL	Δ MEI	D_{OH}	$-\log S_{obs}, m^a$	$-\log S_{calc}, m^b$
1	cyclopentane	18.50	0.3520	0	2.650	2.696
2	methylcyclopentane	22.20	0.4496	0	3.300	3.151
3	2-methylpentane	22.82	0.0782	0	3.790	3.823
4	2,3-dimethylbutane	22.82	0.1662	0	3.610	3.687
5	2-methylhexane	26.52	0.0835	0	4.60	4.420
6	3-methylhexane	26.52	0.1124	0	4.58	4.375
7	2,2-dimethylpentane	26.52	0.2398	0	4.36	4.180
8	2,3-dimethylpentane	26.52	0.2100	0	4.28	4.225
9	3,3-dimethylpentane	26.52	0.2786	0	4.23	4.120
10	3-methylheptane	30.22	0.1210	0	5.16	4.967
11	2,3,4-trimethylpentane	30.22	0.3405	0	4.80	4.630
12	nonane	33.92	0	0	5.88	5.758
13	dodecane	45.02	0	0	7.670	7.572

^a For nos. 1–3, from ref 14, and nos. 4–13, from ref 17. ^b Based on eq 10.

By substituting r_i from eq 2 with r_i from eq 3, we obtain eq 4.

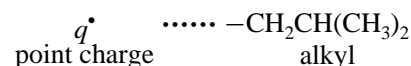
$$E_{(R)} = \frac{-q^2}{2D} \sum \left\{ \alpha_i / l^4 \left[N_i \frac{1 + \cos \theta}{1 - \cos \theta} - \frac{2 \cos \theta (1 - \cos^{N_i} \theta)}{(1 - \cos \theta)^2} \right]^2 \right\} = \frac{-q^2}{2Dl^4} \sum \left\{ \alpha_i \left[N_i \frac{1 + \cos \theta}{1 - \cos \theta} - \frac{2 \cos \theta (1 - \cos^{N_i} \theta)}{(1 - \cos \theta)^2} \right]^2 \right\} \quad (4)$$

For a given system, q , D , and l in eq 4 are all known. If the value α_i is also considered to be constant (2.16 \AA^3), then eq 4 can be further simplified as

$$E_{(R)} = K \sum \left\{ 1 / \left[N_i \frac{1 + \cos \theta}{1 - \cos \theta} - \frac{2 \cos \theta (1 - \cos^{N_i} \theta)}{(1 - \cos \theta)^2} \right]^2 \right\} \quad (5)$$

Here $K = -q^2 \alpha_i / (2Dl^4) = -2.16q^2 / (2Dl^4)$. From eq 5, it can be seen that the relative order of polarizability effect of alkyl substituent will be determined by the right term of $\sum (1/[]^2)$. Therefore, we call it the polarizability effect index (PEI). The PEI value of an alkyl substituent is obtained easily from eq 5, and it quantitates the relative order of the polarizability effect of alkyl substituents. Some PEI values of normal alkyl substituents and increments of PEI (that is, $\Delta PEI = 1/[]^2$) of the i th essential unit from which there are N_i carbon atoms for the point charge (q) are listed in Table 1.

For instance, for the calculation of the PEI for isobutyl, we get



The carbon atom numbers N_i from point charge q to the essential units CH_2- , $CH<$, and CH_3 are 1, 2, and 3, respectively. Thus, the PEI value of i -Bu is $PEI = (1.000\,000 + 0.140\,526 + 2) \times 0.048\,132 = 1.2368$.

In solution, alkane molecules are included by the solvent molecule. Thus, each carbon atom of an alkane molecule will be in the electric field which is formed by solvent molecules and will be polarized. Taking each carbon atom

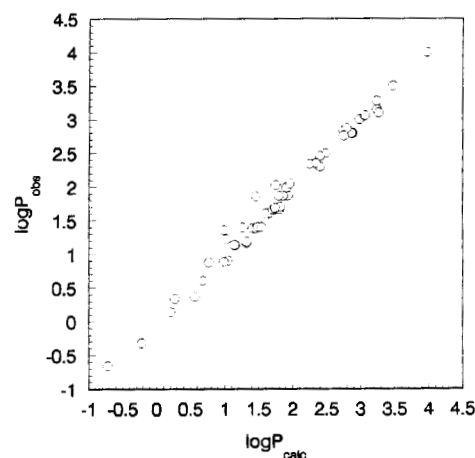
Table 6. SBL, Δ MPEI, D_{OH} , and 1-Octanol–Water Partition Coefficients ($\log P$) for Alcohols and Hydrocarbons

no.	compd	SBL	Δ MPEI	D_{OH}	$\log P_{obs}^a$	$\log P_{calc}^b$
1	methanol	5.6400	0.0000	1.0000	-0.66	-0.72
2	ethanol	9.3400	0.0000	1.0000	-0.32	-0.22
3	propanol	13.0400	0.0000	1.0000	0.34	0.27
4	butanol	16.7400	0.0000	1.0000	0.88	0.77
5	pentanol	20.4400	0.0000	1.0000	1.40	1.26
6	hexanol	24.1400	0.0000	1.0000	2.03	1.75
7	heptanol	27.8400	0.0000	1.0000	2.34	2.25
8	octanol	31.5400	0.0000	1.0000	2.84	2.74
9	nonanol	35.2400	0.0000	1.0000	3.15	3.23
10	2-propanol	13.0400	0.0493	1.0000	0.14	0.21
11	2-methyl-1-propanol	16.7400	0.0687	1.0000	0.61	0.68
12	2-methyl-2-propanol	16.7400	0.1672	1.0000	0.37	0.56
13	3-methyl-1-butanol	20.4400	0.0782	1.0000	1.14	1.16
14	2-methyl-1-butanol	20.4400	0.0976	1.0000	1.14	1.14
15	2-pentanol	20.4400	0.0782	1.0000	1.14	1.16
16	3-pentanol	20.4400	0.0976	1.0000	1.14	1.14
17	3-methyl-2-butanol	20.4400	0.1662	1.0000	0.91	1.05
18	2-methyl-2-butanol	20.4400	0.2155	1.0000	0.89	0.99
19	2,2-dimethyl-1-propanol	20.4400	0.2155	1.0000	1.36	0.99
20	2-hexanol	24.1400	0.0835	1.0000	1.61	1.65
21	3-hexanol	24.1400	0.1123	1.0000	1.61	1.61
22	3-methyl-3-pentanol	24.1400	0.2786	1.0000	1.39	1.41
23	2-methyl-2-pentanol	24.1400	0.2398	1.0000	1.39	1.46
24	2-methyl-3-pentanol	24.1400	0.2100	1.0000	1.41	1.49
25	3-methyl-2-pentanol	24.1400	0.2100	1.0000	1.41	1.49
26	4-methyl-2-pentanol	24.1400	0.1711	1.0000	1.41	1.54
27	2,3-dimethyl-2-butanol	24.1400	0.3473	1.0000	1.17	1.32
28	3,3-dimethyl-1-butanol	24.1400	0.2398	1.0000	1.86	1.46
29	3,3-dimethyl-2-butanol	24.1400	0.3473	1.0000	1.19	1.32
30	2-methyl-2-hexanol	27.8400	0.2537	1.0000	1.87	1.93
31	3-methyl-3-hexanol	27.8400	0.3116	1.0000	1.87	1.86
32	3-ethyl-3-pentanol	27.8400	0.3599	1.0000	1.87	1.80
33	2,3-dimethyl-2-pentanol	27.8400	0.4091	1.0000	1.67	1.74
34	2,3-dimethyl-3-pentanol	27.8400	0.4285	1.0000	1.67	1.71
35	2,4-dimethyl-2-pentanol	27.8400	0.3509	1.0000	1.67	1.81
36	2,4-dimethyl-3-pentanol	27.8400	0.3405	1.0000	1.71	1.82
37	2,2-dimethyl-3-pentanol	27.8400	0.4091	1.0000	1.69	1.74
38	2,2,3-trimethyl-3-pentanol	31.5400	0.6674	1.000	1.99	1.91
39	cyclohexanol	23.5200	0.5127	1.0000	1.23	1.03
40	pentane	19.1200	0.0000	0.0000	2.50	2.48
41	2-methylbutane	19.1200	0.0687	0.0000	2.30	2.40
42	2-methylpentane	22.8200	0.0782	0.0000	2.80	2.88
43	3-methylpentane	22.8200	0.0976	0.0000	2.80	2.86
44	hexane	22.8200	0.0000	0.0000	3.00	2.98
45	heptane	26.5200	0.0000	0.0000	3.50	3.47
46	2,4-dimethylpentane	26.5200	0.1711	0.0000	3.10	3.26
47	octane	30.2200	0.0000	0.0000	4.00	3.97
48	cyclopentane	18.5000	0.3520	0.0000	2.05	1.96
49	cyclohexane	22.2000	0.4003	0.0000	2.46	2.40
50	methylcyclopentane	22.2000	0.4496	0.0000	2.35	2.34
51	cycloheptane	25.9000	0.4828	0.0000	2.87	2.79
52	methylcyclohexane	25.9000	0.5127	0.0000	2.76	2.75
53	cyclooctane	29.6000	0.5253	0.0000	3.28	3.23
54	1,2-dimethylcyclohexane	29.6000	0.6627	0.0000	3.06	3.06

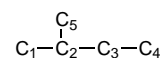
^a From ref 19. ^b Based on eq 12.

as the beginning one in the action of the point charge q , we can calculate its PEI. Then, by summing all of the PEI values, we obtained the MPEI of the alkane molecule in the solvent. That is

$$\text{MPEI} = \sum \text{PEI}(R_i) \quad (6)$$

**Figure 3.** Plot of $\log P_{obs}$ vs $\log P_{calc}$ for alcohols and hydrocarbons.

The calculation of MPEI for 2-methylbutane, for example, is described as follows:



$$\text{C}_1: \text{PEI}(R_1) = 1.2122 + 0.0481 = 1.2603$$

$$\text{C}_2: \text{PEI}(R_2) = 1.887 + 0.1405 \times 2 = 1.4697$$

$$\text{C}_3: \text{PEI}(R_3) = 1.887 + 0.1405 + 0.0481 = 1.3773$$

$$\text{C}_4: \text{PEI}(R_4) = 1.2122 + 0.0235 = 1.2357$$

$$\text{C}_5: \text{PEI}(R_5) = 1.2122 + 0.0481 = 1.2603$$

$$\text{MPEI} = \sum \text{PEI}(R_i) = 1.2603 + 1.4697 + 1.3773 + 1.2357 + 1.2603 = 6.6033$$

That is, the MPEI of 2-methylbutane is 6.6033. Using the same calculation method, we got the MPEI value of n -pentane of 6.5346. Both 2-methylbutane and n -pentane have the same number of carbon atoms, while the MPEI value of 2-methylbutane is more than that of n -pentane. This shows that the former produces a stronger polarizability effect in the electric field than does the latter. Therefore, the MPEI may be used to express the relative order of polarizability effects of alkane molecules in solution. The MPEI values of alkanes bearing 1~10 carbon atoms are listed in Table 2.

3. CORRELATION OF MPEI WITH CAVITY SURFACE AREA (CSA)

In studies of hydrocarbon solubility in water, Hermann¹⁴ developed a method for calculating the surface area inside an idealized solvent cavity required to accommodate the hydrocarbon solute. Kier et al.¹⁵ related the CSA to the molecular connectivity index (χ). In general, the dominant factor of the molecular CSA is the atomic number of type in a molecule. The next factor is the difference in molecular skeleton. For the normal alkane and normal alkyl primary alcohol, there is a linear relationship between CSAs and atomic numbers. Correspondingly, the CSA changes are more complicated and, in general, the CSAs are smaller for branching chain isomers than straight chain isomers. This is caused by the difference in the molecular skeletons. Therefore, we use the sum of bond lengths (SBL) in a molecule to represent the effect of various atoms and atomic numbers on the molecular CSA and use the difference (Δ MPEI, see eq 7) between MPEIs for the branching chain

$$\Delta \text{MPEI} = \text{MPEI}_{br} - \text{MPEI}_{st} \quad (7)$$

and straight chain isomers to represent the effect of the difference in the molecular skeletons on the molecular CSA. In eq 7, the $MPEI_{br}$ and $MPEI_{st}$ are MPEI of branching and straight chain isomers, respectively.

From above, we obtain eq 8. Here, a , b , and c are

$$CSA = a + bSBL + c\Delta MPEI \quad (8)$$

coefficients. The SBL is calculated by using the average values of various chemical bond lengths, that is, the bond lengths of C–C, C–H, C–O, and O–H are 1.54, 1.08, 1.43, and 0.97 (10^{-10}m), respectively. To simplify the calculation procedure, we take the OH group of alcohol as a CH_3 group for computing the $\Delta MPEI$ value. Using the obtained SBL and $\Delta MPEI$ values, we then carried out a correlation of CSA values¹⁵ with the SBL and $\Delta MPEI$ values for the compounds in Table 3 and obtained the regression analysis expression as follows:

$$CSA = 123.626 + 8.65790(SBL) - 99.4269(\Delta MPEI) \quad (9)$$

$$r = 0.9954; \quad s = 5.20; \quad n = 69$$

Kier et al.¹⁵ related the CSA of 69 compounds in Table 3 to the molecular connectivity (χ) and obtained $r = 0.978$ and $s = 11.2$. Equation 9 shows that our result is better than that of the literature.¹⁵ Figure 1 shows a plot of calculated CSA values versus the CSA values in the literature,¹⁵ and the calculated CSAs are in good agreement with the CSAs in the literature.

In eq 9, the coefficient in front of the SBL is positive and that of $\Delta MPEI$ is negative. It shows that the molecular CSA increases as the SBL increases; but, the molecular CSA decreases as the $\Delta MPEI$ of isomers increases. This phenomena may be explained thus: the greater the molecular polarizability effect of isomer, the larger the dipole moment of the isomer introduced by water the larger the dipole moment increases the attraction between the solute and solvent molecule and decreases the distance between the solute and solvent molecule; as a result, the CSA decreases.

Equation 9 also has good prediction ability for CSA. Some predicted values of CSA are listed in Table 4, which are agreement with those of the literature.¹⁴

4. CORRELATION OF MPEI WITH LOG S

It has been shown that the log of the solubilities correlates well with the calculated surface area of the cavity required to accommodate the solute in the solvent water.¹⁴ Therefore, it may be predicted that the solubilities of alkanes and alcohols should also correlate well with the SBL and $\Delta MPEI$ parameters. But, at the same time, attention should be paid to the hydroxyl group OH of alcohol, which is a polar and electron-withdrawing group. Compared with alkanes, the OH group increases the solubility of alcohols. Thus, we take a dummy parameter D_{OH} to represent the contribution of the OH group to the solubility, that is, $D_{OH} = 1$ for the alcohol and $D_{OH} = 0$ for the alkane. Using this treatment, we carried out a regression analysis for the 69 compounds in Table 3 (see also Figure 2) and obtained the correlation of their $-\log S$ values with SBL, $\Delta MPEI$, and D_{OH} :

$$-\log S = 0.2116 + 0.1635(SBL) - 1.535(\Delta MPEI) - 3.079(D_{OH}) \quad (10)$$

$$s = 0.167; \quad R = 0.994; \quad n = 60$$

The solubilities of alkanes predicted by eq 9 are in good agreement with the observed solubilities reported by the literature^{1,14} (see Table 5).

5. CORRELATION OF MPEI WITH LOG P

The partition coefficient of organic compounds of 1-octanol/water is relative to its solubility in water. They have the following linear relationship:¹⁸

$$-\log S = a \log P + b \quad (11)$$

From eq 11, it may be predicted that there is a similar correlation of $\log P$ with SBL, $\Delta MPEI$, and D_{OH} as in eq 10 for alcohol and alkane. We related the $\log P$ of the 54 compounds in Table 6 (see also Figure 3) to the above parameters and obtained the following expression:

$$\log P = -0.0681 + 0.1335(SBL) - 1.241(\Delta MPEI) - 1.401(D_{OH}) \quad (12)$$

$$s = 0.124; \quad R = 0.992; \quad n = 54$$

It can be seen that eq 12 has the same form as eq 10.

6. CONCLUSION

The results of this work show that as the chemical bonding in a molecule increases the CSA is enlarged and the solubilities of alcohols and alkanes are reduced. The molecular polarizability effect also has an influence on CSAs and solubilities. Isomers having larger molecular polarizability effects had smaller CSAs and more solubility. In comparison with alkanes, the OH group of alcohols hardly affects the CSA of an alcohol molecule but enlarges its water solubility much more.

ACKNOWLEDGMENT

Project was supported partly by the NSFC, State Education Commission Foundation (SEC), Ministry of Mechanical Industries Fund (MMF) and Hunan Province Education Commission Fund (HECF) and an award from Hok Yinh-Tong Educational Foundation (HYTF).

REFERENCES AND NOTES

- (1) Katrizky, A. R.; Gordeeva, E. V. Traditional Topological Indices vs Electronic, Geometrical, and Combined Molecular Descriptors in QSAR/QSPR Research. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 835–859.
- (2) Mihalle, Z.; Trinajstić, N. A Graph-Theoretical Approach to Structure-Property Relationships. *J. Chem. Educ.* **1992**, *69*, 701–712.
- (3) Miyashita, Y.; Li, Z.; Sasaki, S. Chemical pattern recognition and multivariate analysis for QSAR studies. *Trends Anal. Chem.* **1993**, *12*, 50–60.
- (4) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17.
- (5) Randić, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6614.

- (6) Balaban, A. T. Highly Discriminating Distance-Based Topological Index. *Chem. Phys. Lett.* **1982**, 89, 399–404.
- (7) Kier, L. B. Indexes of Molecular Shape From Chemical Graphs. *Acta. Pharm. Jugosl.* **1986**, 36, 171–188.
- (8) Osmialowski, K.; Halkiewicz, J.; Radecki, A.; Kaliszan, R. Quantum Chemical Parameters in Correlation Analysis of Gas Liquid Chromatographic Retention Indices of Amines. *J. Chromatogr.* **1985**, 346, 53–60.
- (9) Rolubagh, R.; Jurs, P. C. Descriptions of Molecular Shape Applied in Studies of Structure/Activity and Structure/Property Relationships. *Anal. Chim. Acta* **1987**, 199, 99–109.
- (10) Hall, L. H.; Kier, L. B. Electrotopological State Indices for Atom Types: A Novel Combination of Electronic, Topological, and Valence State Information. *J. Chem. Inf. Comput. Sci.* **1995**, 35, 1039–1045.
- (11) Taft, R. W. The Nature and Analysis of Substituent Electronic Effects. *Prog. Phys. Org. Chem.* **1987**, 16, 1–82.
- (12) Miller, K. J. Additivity Methods in Molecular Polarizability. *J. Am. Chem. Soc.* **1990**, 112, 8533–8542.
- (13) Jenkins, A. D. *Polymer Science*; American Elsevier Publishing Co.: New York, 1972; pp 333.
- (14) Hermann, R. B. Theory of Hydrophobic Bonding. II. The Correlation of Hydrocarbon Solubility in Water with Solvent Cavity Surface Area. *J. Phys. Chem.* **1972**, 76, 2754–2759.
- (15) Kier, L. B.; Hall, L. H.; Murray, W. J.; Randic, M. Molecular Connectivity I: Relationship to Nonspecific Local Anesthesia. *J. Pharm. Sci.* **1975**, 64, 1971–1974.
- (16) Hall, L. H.; Kier, L. B.; Murray, W. J. Molecular Connectivity II: Relationship to Water Solubility and Boiling Point. *J. Pharm. Sci.* **1975**, 64, 1974–1977.
- (17) Nelson, T. M.; Jurs, P. C. Prediction of Aqueous Solubility of Organic Compounds. *J. Chem. Inf. Sci.* **1994**, 34, 601–609.
- (18) Hansch, C.; Quinlan, J. E.; Lawrence, G. L. The Linear Free-Energy Relationship between Partition Coefficients and the Aqueous Solubility of Organic Liquids. *J. Org. Chem.* **1968**, 33, 347–350.
- (19) Murray, W. J.; Hall, L. H.; Kier, L. B. Molecular Connectivity III: Relationship to Partition Coefficients. *J. Pharm. Sci.* **1975**, 64, 1978–1981.

CI9601729