QSPR Modeling for Solubility of Fullerene (C$_{60}$) in Organic Solvents[†]

N. Sivaraman, T. G. Srinivasan, and P. R. Vasudeva Rao*

Fuel Chemistry Division, Chemical Group, Indira Gandhi Centre for Atomic Research,
Kalpakkam - 603 102, India

R. Natarajan

Campion Higher Secondary School, Tiruchirappalli - 620 001, India

Received February 22, 2001

Solubility of fullerene C$_{60}$ in 75 organic solvents was examined to develop quantitative structure-solubility relationships. Topological indices and polarizability parameter computed from refractive index were used to form the regression models. The models suggested for individual data sets such as alkanes, aklyl halides, alcohols, cycloalkanes, alkylbenzenes, and aryl halides have good predictive ability and are better than the models for the combined groups. Inclusion of an indicator parameter which is a combination of atom contributions and contributions of substituents' position in benzenes improved the predictive ability significantly.

INTRODUCTION

Quantitative information on the solubility of C$_{60}$ in various organic solvents was reported from our laboratory earlier.[1−3] Other research groups reported the solubility of C$_{60}$ include Ruoff et al.,[4] Scrivens et al.,[5,6] and Beck et al.[7−8] Solubility of C$_{60}$ reported by various research groups were consolidated and reviewed by Beck and Mandi.[9]

In our earlier work,[3] we examined the solubility of C$_{60}$ in about 30 solvents and studied the influence of solvent properties such as polarizability parameter, solubility parameter, molar volume, etc. Ruoff et al.[4] also made an attempt to correlate C$_{60}$ solubility to various solvent properties. The above studies indicated only certain trends in solubility; however, there has been no systematic study on the correlation of C$_{60}$ solubility and solvent properties. Moreover, with the availability of solubility data now for C$_{60}$ in about 100 solvents, a better appreciation of solubility behavior could be achieved. The solubility model can be used for predicting the solubility of C$_{60}$ in solvents for which experimental data are not available. Methods to predict solubility of C$_{60}$ in organic solvents can also help the fullerene researchers to select solvents for its purification and chromatographic separations. QSPR approach has been found to be very useful in developing predictive models.

Kamlet and co-workers modeled solubility[10] and related phenomena using a cavity term, a dipolar term, and terms of H-bonding. This approach known as the Solvatochromic approach formed the basis for the evolution of other models called the Linear Solvation Energy Relationships (LSER). These models were proposed for the estimation of aqueous solubility of the organic compounds and reviewed by Reinhard and Axel Drefahl.[11]

In the case of aqueous solubility of the organic compounds, the solute molecules are different in each case, but the solvent remains the same. In the solubility of C$_{60}$ in organic solvents, the situation is different because the solute is the same but the solvents are different. Hence, cavity surface area used in the previous study[12] cannot be used here. The present study being the first attempt in the QSAR/QPPR modeling of solubility of C$_{60}$, we preferred the use of conceptually simple, fundamental, and easily computable properties or parameters.

In our approach, we used polarizability parameter and the connectivity indices to form the regression models. Polarizability parameter can be easily calculated from the refractive index. It is well established that topological parameters such as Randic connectivity indices, Hall and Kier valance connectivity indices, and other graph theoretical indices encode the structural details of compounds.[13] Connectivity indices in the regression model can explain the type of structural feature that affects the solubility of fullerene in the organic solvents.

EXPERIMENTAL SECTION

Solubility Data. The experimental details regarding solubility of C$_{60}$ in various organic solvents have been reported in detail.[1−3] The solubility of C$_{60}$ in acetone, methanol, acetonitrile, and tetrahydrofuran are below detectable limits (less than 1 μg/mL). Hence, these solvents were not included in the regression analysis. A few more solvents such as benzonitrile, nitro compounds, substituted naphthalenes, and alkenes were not included in the present study as there are only one or two observations in each of their class. In the construction of regression models, log(1/S) values were used (S is solubility in mol/L). The solubility data set was divided into subsets based on the chemical nature of the solvents. Data subsets and the number of solvents are given in Table 1. The data sets were used separately and in combination. The topological indices used are listed in Table 2 with their symbols. The summary of regression analysis

Table 1. Division of Dataset into Subsets

subset	class/type of solvents	no. of observations
A	alkanes	7
B	alkyl halides	32
C	alcohols	6
D	cycloalkanes	5
E	alkylbenzenes	16
F	aryl halides	9

Table 2. Symbols and Definitions of Parameters Used in the Regression Models

symbol	definition
$^h\chi$	connectivity index of order $h = 0-6$
$^h\chi_c$	cluster connectivity index of order $h = 3-6$
$^h\chi_{pc}$	path-cluster connectivity index of order $h = 4-6$
$^h\chi^v$	valence connectivity index of order $h = 0-6$
$^h\chi^v_c$	valence cluster connectivity index of order $h = 3-6$
$^h\chi^v_{pc}$	valence path-cluster connectivity index of order $h = 4-6$
$^h\chi^v_{ch}$	valence chain connectivity index of order $h = 6$
F	number of fluorine atoms
Cl	number of chlorine atoms
Br	number of bromine atoms
I	number of iodine atoms
OH	number of alcohol groups
o	indicator parameter for ortho substituents
m	indicator parameter for meta substituents
p	indicator parameter for para substituents
Φ	polarizability parameter

is given in Table 3. The solubility data used in the present study include the experimental data of the authors[3] along with the other data reported[4-9] and are presented in Tables 4-7. The topological indices used as independent parameters are given in Table 4 (for aliphatic compounds) and Table 6 (for aromatic compounds).

Polarizability parameter Φ was computed for each solvent from the refractive index using eq 1.

$$\Phi = (n^2 - 1)/(n^2 + 2) \tag{1}$$

Randic connectivity indices and Hall and Kier connectivity indices for aliphatic compounds were calculated based on their definition.[13a] A topological calculator that is being developed by one of the authors[14] was used to compute the connectivity indices of aliphatic compounds. In the case of benzene and alkyl benzenes, they were taken from the book of Hall and Kier.[15] For aryl halides, they were calculated using the group exchange algorithm explained by Hall and Kier.[16]

In addition to these parameters, an indicator parameter is used. The indicator parameter IP is a combination of atom contributions, group contribution, and contribution due to the position of an atom or a group in aromatic compounds. IP is given by the relation

$$IP = a(F) + b(Cl) + c(Br) + d(I) + e(OH) +$$
$$f(\text{ortho sub}) + g(\text{meta sub}) + h(\text{para sub}) \tag{2}$$

where the coefficients in lower cases are obtained from regression analysis. For a given solvent, all the indicators or contribution parameters were not used. This approach of including group and atom contribution parameters had been used in other situations[17] and is not new in QSAR analysis.

Statistical analyses are carried out using SPSS 10.0 for Windows.

Table 3. Summary of Regression Analysis

regression analysis	parameters used	r	r^2	SE
	Set A /Alkanes (Seven Solvents)			
using ϕ only	Φ	0.926	0.857	0.207
using χ only	$^1\chi$	0.846	0.716	0.292
using ϕ and χ	$\Phi, ^2\chi$	0.953	0.907	0.186
	Set A after Deleting Hexane (Six Solvents)			
using ϕ only	Φ	0.995	0.990	0.0626
using χ only	$^0\chi$	0.955	0.912	0.181
using ϕ and χ	$\Phi, ^3\chi_c$	0.997	0.993	0.0576
	Set B/Alkylhalides (32 Solvents)			
using ϕ only	Φ	0.860	0.740	0.367
using ϕ and IP	Φ, IP	0.943	0.889	0.258
using χ only	$^1\chi$	0.680	0.463	0.537
using χ only	$^1\chi_v, ^2\chi, ^3\chi^v_c, ^4\chi^v_{pc}$	0.788	0.621	0.468
using χ and IP	$^1\chi_v, ^2\chi, ^3\chi_c, ^3\chi_cv, ^4\chi_{pc}v$, IP	0.964	0.929	0.225
using ϕ and χ	$\Phi, ^1\chi, ^1\chi^v, 2\chi, ^3\chi^v_c, ^4\chi^v_{pc}$	0.942	0.887	0.265
using ϕ, χ, and IP	$\Phi, ^1\chi, ^1\chi_v, 2\chi, ^3\chi^v_c, ^4\chi^v_{pc}$, IP	0.967	0.935	0.220
	Set C/Alcohols (Six Solvents)			
using ϕ only	Φ	0.986	0.973	0.123
using χ only	$^3\chi$	0.942	0.887	0.250
using ϕ and χ	$\Phi, ^3\chi$	0.993	0.986	0.0997
	Set D/Cycloalkanes (Five Solvents)			
using ϕ only	Φ	0.890	0.792	0.447
using χ only	$^5\chi$	0.989	0.978	0.146
using ϕ and χ	$\Phi, ^5\chi$	0.999	0.999	0.0428
	Set E/Alkylbenzenes (16 Solvents)			
using ϕ only	Φ	0.709	0.503	0.310
using ϕ and IP	Φ, IP	0.918	0.842	0.197
using χ only	$^4\chi^v_p, ^6\chi^v_{ch}$	0.795	0.632	0.277
using χ and IP	$^4\chi^v_p, ^6\chi^v_{ch}$, IP	0.960	0.922	0.145
using ϕ and IP	$^6\chi^v_{pc}$, IP	0.974	0.949	0.112
	Set F/Aryl Halides (Nine Solvents)			
using ϕ only	Φ	0.526	0.277	0.466
using ϕ and IP	Φ, IP	0.801	0.641	0.435
using χ only	$^1\chi^v$	0.647	0.417	0.418
using χ and IP	$^6\chi^v_p$, IP	0.994	0.988	0.157
	Combined Set AB/Aliphatic (39 Solvents)			
using ϕ and χ	$\Phi, ^4\chi^v_{pc}, ^3\chi^v_c, ^1\chi, ^3\chi, ^0\chi^v$	0.952	0.906	0.241
using ϕ, χ, and IP	$\Phi, ^4\chi^v_{pc}, ^3\chi^v_c, ^1\chi, ^3\chi, ^0\chi^v$, IP	0.960	0.922	0.233
	Combined Set ABC/Aliphatic (45 Solvents)			
using ϕ and IP	Φ, IP	0.954	0.910	0.256
using ϕ, χ, and IP	$\Phi, ^4\chi^v_{pc}, ^2\chi$, IP	0.958	0.917	0.253
	Combined Set EF/Aromatics (25 Solvents)			
using χ and IP	$^6\chi^v_{pc}$, IP	0.924	0.854	0.217
using χ and IP	$^0\chi^v, ^1\chi^v, ^2\chi^v, ^6\chi^v_p$, IP	0.971	0.943	0.207
	Validation Study/Aliphatic (Training 29 + Testing 16 Solvents)			
using ϕ, χ, and IP	$\Phi, ^4\chi^v_{pc}$, IP	0.965		0.254

RESULTS AND DISCUSSION

The applicability of polarizability parameter Φ alone was first examined in the data sets and the ANOVA results are given in Table 3. Regression analyses were rerun after including the IP value which has the contribution for the various atoms, groups, and the pattern of substitution. The minimum number of chi (χ) values required for regression model with a possible low standard error of the estimation was selected for each model by the backward multiple regression analysis procedure. Using these χ values and the IP values, stepwise multiple regression procedure was followed to determine the coefficients of the parameters in the expression for IP. To test whether there was any decrease in the standard error, Φ value was included, and the multiple regression procedure was repeated. The results thus obtained for each data set and the combined set are given in Table 3.

Table 4. Experimental Solubility of C$_{60}$ in Aliphatic Compounds and Cycloalkanes with Independent Parameters

no.	solvent	log(1/S)	Φ	$^0\chi^v$	$^1\chi^v$	$^3\chi^v_c$	$^4\chi^v_{pc}$	$^1\chi$	$^2\chi$
1	pentane	5.255	0.220	4.121	2.414	0.000	0.000	2.414	1.354
2	hexane	4.255	0.232	4.828	2.914	0.000	0.000	2.914	1.707
3	octane	4.459	0.238	6.243	3.914	0.000	0.000	3.914	2.457
4	isooctane	4.442	0.241	6.784	3.417	1.969	0.816	3.417	4.159
5	decane	4.012	0.248	7.656	4.914	0.000	0.000	4.914	3.122
6	dodecane	3.898	0.254	9.069	5.914	0.000	0.000	5.914	3.829
7	tetradecane	3.757	0.257	10.484	6.914	0.000	0.000	6.914	4.536
8	iodomethane	2.971	0.310	3.500	2.500	0.000	0.000	1.000	0.000
9	bromoethane	4.012	0.255	3.668	2.093	0.000	0.000	1.414	0.707
10	iodoethane	3.410	0.301	4.207	2.475	0.000	0.000	1.414	0.707
11	1-choropropane	4.556	0.236	3.546	2.007	0.000	0.000	1.914	1.000
12	1-bromopropane	4.158	0.260	4.375	2.594	0.000	0.000	1.914	1.000
13	1-iodopropane	3.627	0.296	4.914	2.974	0.000	0.000	1.914	1.000
14	2-chloropropane	4.857	0.231	3.709	1.807	0.653	0.000	1.732	1.732
15	2-bromopropane	4.380	0.256	4.538	2.285	1.131	0.000	1.732	1.732
16	2-iodopropane	3.816	0.293	5.077	2.597	1.443	0.000	1.732	1.732
17	1-choloro-2-methylpropane	4.380	0.241	4.416	2.362	0.408	0.462	2.270	1.802
18	1-bromo-2-methylpropane	3.903	0.261	5.245	2.948	0.408	0.800	2.270	1.802
19	1-iodo-2-methylpropane	3.326	0.292	5.784	3.329	0.408	1.020	2.270	1.802
20	2-chloro-2-methypropane	4.857	0.234	4.632	2.066	2.132	0.000	2.000	3.000
21	2-bromo-2-methylpropane	4.079	0.257	5.461	2.480	2.961	0.000	2.000	3.000
22	2-iodo-2-methylpropane	3.496	0.289	6.000	2.750	3.500	0.000	2.000	3.000
23	dibromomethane	3.301	0.314	4.629	2.773	0.000	0.000	1.414	0.707
24	bromochloromethane	2.982	0.286	3.800	2.187	0.000	0.000	1.414	0.707
25	1,2-dichloroethane	3.954	0.266	3.678	2.100	0.000	0.000	1.914	1.000
26	1,2-dibromoethane	3.158	0.313	5.336	3.273	0.000	0.000	1.914	1.000
27	1,2-dichloropropane	3.857	0.263	4.548	2.438	0.462	0.523	2.270	1.802
28	1,2-dibromopropane	3.313	0.304	6.206	3.503	0.800	1.569	2.270	1.802
29	1,3-dichloropropane	3.778	0.268	4.385	2.600	0.000	0.000	2.414	1.353
30	1,3-dibromopropane	3.255	0.305	6.043	3.773	0.000	0.000	2.414	1.353
31	1,3-diiodopropane	2.415	0.361	7.121	4.535	0.000	0.000	2.414	1.353
32	dichloromethane	3.459	0.255	2.971	1.601	0.000	0.000	1.414	0.707
33	chloroform	3.653	0.267	3.973	1.959	0.837	0.000	1.732	1.732
34	bromoform	2.106	0.340	6.460	3.394	4.351	0.000	1.732	1.732
35	1,1,1-trichloroethane	3.681	0.262	4.896	2.198	2.732	0.000	2.000	3.000
36	1,2,3-trichloropropane	3.051	0.285	5.387	3.069	0.327	0.740	2.808	1.922
37	carbontetrachloride	3.204	0.274	5.028	2.264	2.903	0.000	2.000	3.000
38	1,1,2-trichlorotrifluoroethane	4.857	0.220	5.530	2.515	1.107	1.668	3.250	4.500
39	tetrachloroethane	2.133	0.291	5.682	2.945	0.855	1.932	2.643	2.488
40	ethanol	5.857	0.221	2.154	1.023	0.000	0.000	1.414	0.707
41	1-propanol	5.244	0.234	2.861	1.523	0.000	0.000	1.914	1.000
42	1-butanol	4.884	0.242	3.568	2.023	0.000	0.000	2.414	1.354
43	1-penatnol	4.380	0.247	4.275	2.523	0.000	0.000	2.914	1.707
44	1-hexanol	4.234	0.252	4.982	3.023	0.000	0.000	3.414	2.061
45	1-octanol	4.185	0.258	6.396	4.023	0.000	0.000	4.414	2.768
46	cyclopentane	5.556	0.245						
47	cyclohexane	4.150	0.256						
48	methylcyclohexane	3.627	0.254						
49	1,2-dimethylcyclohexane	3.743	0.261						
50	ethylcyclohexane	3.459	0.259						

In this table, r^2 denotes correlation coefficient and "SE" denotes standard error of estimate. The best model in each set and the combined sets are discussed below.

Study of Set A — Alkanes. Polarizibility parameter Φ explained 85.7% of the data variability. When hexane alone was deleted from the data list, Φ alone explained 99.0% of the data variability, and the standard error was only 0.0626. When $^3\chi_c$ was included, the standard error dropped to 0.0576. This was taken as the best regression model for this set and the equation is

$$\log(1/S) = 13.958 - 39.762(\Phi) + 0.0422(^3\chi_c) \quad (3)$$

$$n = 6, r = 0.997, r^2 = 0.993, \text{ and SE} = 0.0576$$

where n, r, and SE are number of data, correlation coefficient, and standard error of the estimate, respectively.

The predicted solubility values using the above model are given in Table 5 (solvents 1−7, under the column "Pred^a").

Study of Set B — Alkyl Halides. This set consisted of 32 alkyl halides and Φ alone could explain only 74.0% of the data variability. This showed that in halogenated solvents, the solute−solvent interaction cannot be explained fully by polarizability. The number and the nature of each halogen atom seemed to influence the solubility. This was very much reflected by the fact that there was an appreciable improvement in the correlation coefficient and decrease in the standard error of the estimate when IP value was introduced into the regression model. The four χ values selected by the backward multiple regression procedure gave a standard error of 0.468, and when IP was included, it decreased to 0.225. When Φ was also included, there was only a slight improvement and the standard error became 0.220. This

Table 5. Experimental and Predicted Solubility of C_{60} in Aliphatic Compounds and Cycloalkanes

no.	solvent	solubility in log(1/S)						solubility in mg/mL					
		exptl	Pred[a]	AB1[b]	AB2[c]	ABC1[d]	ABC2[e]	exptl	pred[f]	AB1[g]	AB2[h]	ABC1[i]	ABC2[j]
1	pentane	5.255	5.226	5.010	4.937	5.079	5.102	0.004	0.004	0.007	0.008	0.006	0.0057
2	hexane	4.255	4.749	4.599	4.605	4.674	4.703	0.040	0.013	0.018	0.018	0.015	0.0143
3	octane	4.459	4.490	4.396	4.397	4.455	4.466	0.025	0.023	0.029	0.029	0.025	0.0246
4	isooctane	4.442	4.442	4.282	4.485	4.343	4.192	0.026	0.026	0.038	0.024	0.033	0.0463
5	decane	4.012	4.085	4.061	4.091	4.111	4.113	0.070	0.059	0.063	0.058	0.056	0.0555
6	dodecane	3.898	3.854	3.881	3.900	3.915	3.900	0.091	0.101	0.095	0.091	0.088	0.0906
7	tetradecane	3.757	3.727	3.795	3.778	3.807	3.771	0.126	0.135	0.115	0.120	0.112	0.1219
8	iodomethane	2.971	2.945	2.875	3.121	3.031	3.123	0.770	0.818	0.960	0.545	0.670	0.5426
9	bromoethane	4.012	4.166	4.381	4.305	4.357	4.398	0.070	0.049	0.030	0.036	0.032	0.0288
10	iodoethane	3.410	3.406	3.193	3.373	3.345	3.397	0.280	0.282	0.461	0.305	0.326	0.2889
11	1-choropropane	4.556	4.435	4.546	4.468	4.526	4.579	0.020	0.026	0.020	0.025	0.021	0.0190
12	1-bromopropane	4.158	4.174	4.348	4.236	4.178	4.216	0.050	0.048	0.032	0.042	0.048	0.0438
13	1-iodopropane	3.627	3.554	3.507	3.558	3.496	3.531	0.170	0.201	0.224	0.199	0.230	0.2121
14	2-chloropropane	4.857	4.650	4.700	4.683	4.712	4.730	0.010	0.016	0.014	0.015	0.014	0.0134
15	2-bromopropane	4.380	4.301	4.365	4.325	4.330	4.334	0.030	0.036	0.031	0.034	0.034	0.0333
16	2-iodopropane	3.816	3.609	3.401	3.541	3.594	3.597	0.110	0.177	0.286	0.207	0.183	0.1820
17	1-choloro-2-methylpropane	4.380	4.378	4.404	4.365	4.354	4.326	0.030	0.030	0.028	0.031	0.032	0.0340
18	1-bromo-2-methylpropane	3.903	4.056	4.176	4.086	4.155	4.061	0.090	0.063	0.048	0.059	0.050	0.0626
19	1-iodo-2-methylpropane	3.326	3.431	3.401	3.442	3.652	3.518	0.340	0.267	0.286	0.260	0.161	0.2186
20	2-chloro-2-methylpropane	4.857	4.734	4.618	4.666	4.587	4.564	0.010	0.013	0.017	0.016	0.019	0.0196
21	2-bromo-2-methylpropane	4.079	4.359	4.287	4.293	4.280	4.239	0.060	0.032	0.037	0.037	0.038	0.0415
22	2-iodo-2-methylpropane	3.496	3.716	3.487	3.618	3.746	3.696	0.230	0.139	0.235	0.174	0.129	0.1451
23	dibromomethane	3.301	3.103	3.047	3.083	2.826	2.896	0.360	0.568	0.646	0.595	1.074	0.9146
24	bromochloromethane	2.982	3.426	3.401	3.429	3.322	3.401	0.750	0.270	0.286	0.268	0.343	0.2859
25	1,2-dichloroethane	3.954	3.712	3.609	3.623	3.532	3.622	0.080	0.140	0.177	0.171	0.212	0.1720
26	1,2-dibromoethane	3.158	3.203	3.244	3.183	2.867	2.924	0.500	0.451	0.410	0.473	0.978	0.8576
27	1,2-dichloropropane	3.857	3.738	3.701	3.686	3.626	3.615	0.100	0.132	0.143	0.148	0.170	0.1748
28	1,2-dibromopropane	3.313	2.896	2.927	2.850	3.187	2.998	0.350	0.914	0.852	1.018	0.468	0.7237
29	1,3-dichloropropane	3.778	3.783	3.516	3.529	3.454	3.535	0.120	0.119	0.219	0.213	0.253	0.2102
30	1,3-dibromopropane	3.255	3.419	3.508	3.351	3.130	3.162	0.400	0.275	0.224	0.321	0.534	0.4954
31	1,3-diiodopropane	2.415	2.407	2.390	2.412	2.306	2.307	2.770	2.822	2.931	2.790	3.563	3.5512
32	dichloromethane	3.459	3.775	3.818	3.823	3.879	3.964	0.250	0.121	0.109	0.108	0.095	0.0782
33	chloroform	3.653	3.499	3.591	3.548	3.492	3.549	0.160	0.228	0.185	0.204	0.232	0.2035
34	bromoform	2.106	2.055	2.217	2.094	2.415	2.427	5.641	6.338	4.368	5.794	2.767	2.6966
35	1,1,1-trichloroethane	3.681	3.625	3.725	3.669	3.657	3.660	0.150	0.171	0.136	0.154	0.159	0.1575
36	1,2,3-trichloropropane	3.051	3.083	3.066	2.992	2.868	2.850	0.640	0.594	0.618	0.734	0.976	1.0166
37	carbontetrachloride	3.204	3.121	3.366	3.242	3.250	3.263	0.450	0.545	0.310	0.413	0.405	0.3931
38	1,1,2-trichlorotrifluoroethane	4.857	4.857	4.702	4.857	4.857	4.857	0.010	0.010	0.014	0.010	0.010	0.0100
39	tetrachloroethane	2.133	2.373	2.526	2.427	2.676	2.484	5.301	3.051	2.144	2.694	1.517	2.3614
40	ethanol	5.857	5.892			5.522	5.522	0.001	0.001			0.002	0.0022
41	1-propanol	5.244	5.229			5.087	5.096	0.004	0.004			0.006	0.0058
42	1-butanol	4.884	4.752			4.817	4.825	0.009	0.013			0.011	0.0108
43	1-pentanol	4.380	4.470			4.622	4.625	0.030	0.024			0.017	0.0171
44	1-hexanol	4.234	4.280			4.466	4.464	0.042	0.038			0.025	0.0247
45	1-octanol	4.185	4.161			4.271	4.252	0.047	0.050			0.039	0.0403
46	cyclopentane	5.556	5.540					0.002	0.002				
47	cyclohexane	4.150	4.193					0.051	0.046				
48	methylcyclohexane	3.627	3.642					0.170	0.164				
49	1,2-dimethylcyclohexane	3.743	3.736					0.130	0.132				
50	ethylcyclohexane	3.459	3.424					0.250	0.271				

[a] The values are calculated for individual sets, solvents 1−7 using eq 3, solvents 8−39 using eq 4, solvents 40−45 using eq 5, and solvents 46−50 using eq 6. [b] Predicted using eq 9 for combined Set AB. [c] Predicted using eq 10 for combined Set AB. [d] Predicted using eq 11 for combined Set ABC. [e] Predicted using eq 12 for combined Set ABC. [f]−[j] Solubility in mg/mL obtained from [a]−[e], respectively.

regression equation is given below and is used to predict the solubility (Table 5, solvents 8−39, under the column "Pred[a]") for this data set

$$\log(1/S) = 7.828 - 14.264(\Phi) - 0.204(^3\chi^v_c) - 0.367(^4\chi^v_{pc}) + 0.353(^2\chi) - 0.219(^1\chi^v) + 0.178(^1\chi) + IP \quad (4)$$

where $IP = 0.07512(F) - 0.282(Cl) - 0.06912(Br) - 0.09373(I)$, $n = 32$, $r = 0.967$, $r^2 = 0.935$, and $SE = 0.220$.

The above model has more independent parameters than usual. When the two-parameter model involving Φ and IP

with a standard error of 0.258 was used to predict the $\log 1/S$ values and on conversion to solubility in mg/mL, a large variation between experimental and predicted values was observed. To have a reasonably good correlation between the experimental and the predicted solubility in mg/mL, the standard error of the estimate has to be very low. To achieve this, we have to use more parameters than usual as in the above model. Considering the good predictive ability of the model, this may be tolerated.

Study of Set C − Alcohols. For alcohols, Φ alone explained 97.3% of data variability. Hence, it may be concluded that in the case of alcohols in which C_{60} has low

Table 6. Solubility of C_{60} in Aromatic Compounds with Independent Parameters used in the Regression Models

no.	solvent	log(1/S)	Φ	$^0\chi^v$	$^1\chi^v$	$^2\chi^v$	$^6\chi^v_p$	$^6\chi^v_{pc}$
1	benzene	2.629	0.295	3.464	2.000	1.155	0.000	0.000
2	toluene	2.410	0.293	4.387	2.411	1.655	0.064	0.128
3	1,2-dimethylbenzene	1.930	0.297	5.309	2.827	2.084	0.111	0.415
4	1,3-dimethylbenzene	2.681	0.293	5.309	2.821	2.158	0.159	0.367
5	1,4-dimethylbenzene	2.094	0.292	5.309	2.821	2.155	0.111	0.415
6	1,3,5-trimethylbenzene	2.687	0.293	6.232	3.232	2.665	0.269	0.683
7	1,2,3-trimethylbenzene	2.185	0.302	6.232	3.244	2.517	0.192	0.904
8	1,2,4-trimethylbenzene	1.605	0.296	6.232	3.238	2.588	0.186	0.839
9	1,2,3,4-tetramethylbenzene	2.101	0.303	7.155	3.661	2.949	0.250	1.577
10	1,2,3,5-tetramethylbenzene	1.539	0.300	7.155	3.655	3.024	0.281	1.441
11	1-ethylbenzene	2.442	0.292	5.094	2.971	1.839	0.124	0.209
12	1-propylbenzene	2.681	0.290	5.801	3.471	2.236	0.197	0.270
13	isopropylbenzene	2.778	0.290	5.964	3.354	2.565	0.165	0.489
14	1-butylbenzene	2.579	0.289	6.508	3.971	2.589	0.351	0.284
15	*sec*-butylbenzene	2.816	0.289	6.671	3.892	2.720	0.225	0.678
16	*tert*-butylbenzene	2.903	0.290	6.887	3.661	3.616	0.199	1.058
17	fluorobenzene	3.086	0.277	3.765	2.100	1.296	0.024	0.049
18	chlorobenzene	2.018	0.306	4.519	2.477	1.731	0.073	0.145
19	bromobenzene	2.339	0.323	5.348	2.891	2.210	0.126	0.252
20	iodobenzene	2.535	0.351	5.887	3.161	2.520	0.160	0.328
21	1,2-dichlorobenzene	1.442	0.319	5.574	2.960	2.227	0.126	0.498
22	1,2-dibromobenzene	1.717	0.347	7.232	3.789	3.120	0.218	1.176
23	1,3-dichlorobenzene	2.477	0.317	5.574	2.954	2.311	0.188	0.437
24	1,3-dibromobenzene	1.717	0.346	7.232	3.783	3.268	0.403	0.991
25	1,2,4-trichlorobenzene	1.928	0.329	6.629	3.436	2.807	0.163	1.019

solubility, polarizability is able to encode the intermolecular forces responsible for dissolution. Structural parameter in terms of $^3\chi$ is also able to explain 88.7% of data variability. A combination of these two terms gives the best regression model with the standard error of 0.0997. The predicted solubility is given in Table 5 (solvents 40−45, under the column "Pred^a"). The regression equation is

$$log(1/S) = 22.792 - 76.572(\Phi) + 0.650(^3\chi) \quad (5)$$

$$n = 6, \ r = 0.993, \ r^2 = 0.986, \ \text{and SE} = 0.0997$$

Study of Set D − Cycloalkanes. The solvents in this class were cycloalkanes. Being cyclic in nature, the structural parameter was able to explain a higher percentage of data variability. When the solvent molecule is cyclic, structure-solubility relationship seems to be dominant. Inclusion of both the Φ and χ term ($^5\chi$) accounted for 99.9% of data variability. The regression model for this set of solvents is

$$log(1/S) = 61.078(\Phi) - 2.716(^5\chi) - 9.418 \quad (6)$$

$$n = 5, \ r = 0.999, \ r^2 = 0.999, \ \text{and SE} = 0.0428$$

The predicted solubility is shown in Table 5 (solvents 46−50, under the column "Pred^a").

Study of Set E − Alkyl Benzenes. The solvents in this class were aromatic hydrocarbons. As mentioned in Set D, in this set also Φ had very poor correlation with log(1/S) and could explain only 50.3% of data variability. When IP values for the pattern of substitution were included, the standard error decreased from 0.310 to 0.197. This explained that the position of substitution played an important role in deciding the solubility of C_{60} in aromatic hydrocarbons. When χ terms were used to form the regression model, the higher order χ terms namely, $^4\chi^v_p$ and $^6\chi^v_{ch}$, explained solubility better than the other terms. A combination of $^6\chi^v_{pc}$ and IP values formed the best regression model with good

predictive ability (SE = 0.112). Inclusion of Φ had not improved the correlation. The regression equation for this set is

$$log(1/S) = 2.444 + 0.507(^6\chi^v_{pc}) + IP \quad (7)$$

where IP = $-0.620(o)$ $-0.03604(m)$ $-$ $0.522(p)$, $n = 16$, $r = 0.974$, $r^2 = 0.949$, and SE = 0.112.

The correlation of higher order terms with the solubility indicates that in the case of aromatic compounds, solubility of C_{60} depends on the structure of the solvent molecule. This observation is analogous to the situation in cycloalkanes. Predicted solubility for this set is given in Table 7 (solvents 1−16, under the column "Pred^a") along with experimental solubility.

Study of Set F − Halobenzenes. In the case of aryl halides, polarizibility parameter was very poorly correlated with the solubility. It had already been noted in alkyl halides that when halogens were present, Φ was not a very good solvent parameter to explain the solubility. In addition to this, the compounds in the set being cyclic, poor correlation of Φ with solubility was an expected phenomenon. There was no appreciable change when the number of halogen atoms and the indicators for substitution were included. On the other hand, structural parameters gave good correlation. The best regression model for the set was obtained by using $^6\chi^v_p$ and indicator parameters. The regression equation is

$$log(1/S) = 3.373 - 0.9023(^6\chi^v_p) + IP \quad (8)$$

where IP = $0.356(o) + 1.986(m) + 1.467(p) - 0.06752(F)$ $- 0.598(Cl) + 0.610(I)$, $n = 9$, $r = 0.994$, $r^2 = 0.988$, and SE = 0.157.

Again a higher order χ term explains the solubility. In the case of cyclic compounds, the solubility of C_{60} was related to higher order χ terms. Hence, it may be concluded that the solubility of C_{60} in a cyclic solvent depends very much

Table 7. Experimental and Predicted Solubility of C_{60} in Aromatic Compounds

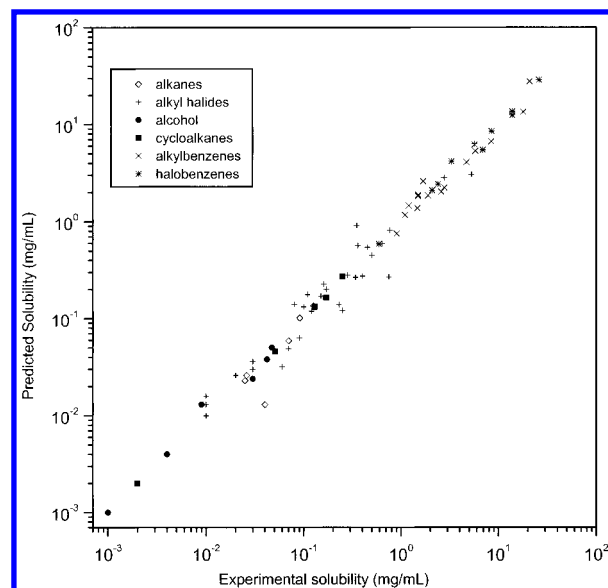|		log(1/S)				solubility in mg/mL			
no.	solvent	exptl	Pred[a]	EF1[b]	EF1[c]	exptl	Pred[d]	EF1[e]	EF2[f]
1	benzene	2.629	2.444	2.392	2.577	1.690	2.592	1.906	2.922
2	toluene	2.410	2.509	2.456	2.245	2.800	2.231	4.092	2.521
3	1,2-dimethylbenzene	1.930	2.034	1.961	1.946	8.459	6.659	8.151	7.872
4	1,3-dimethylbenzene	2.681	2.594	2.575	2.549	1.500	1.835	2.035	1.918
5	1,4-dimethylbenzene	2.094	2.133	2.227	2.259	5.800	5.307	3.967	4.268
6	1,3,5-trimethylbenzene	2.687	2.718	2.732	2.782	1.479	1.378	1.190	1.333
7	1,2,3-trimethylbenzene	2.185	2.246	2.205	2.229	4.700	4.084	4.254	4.487
8	1,2,4-trimethylbenzene	1.605	1.727	1.801	1.825	17.899	13.488	10.768	11.380
9	1,2,3,4-tetramethylbenzene	2.101	2.066	2.170	2.105	5.701	6.181	5.654	4.870
10	1,2,3,5-tetramethylbenzene	1.539	1.413	1.465	1.299	20.799	27.818	36.210	24.708
11	1-ethylbenzene	2.442	2.550	2.496	2.482	2.600	2.032	2.374	2.299
12	1-propylbenzene	2.681	2.581	2.527	2.769	1.500	1.890	1.226	2.142
13	Isopropylbenzene	2.778	2.692	2.635	2.642	1.200	1.464	1.641	1.667
14	1-butylbenzene	2.579	2.588	2.533	2.589	1.900	1.861	1.855	2.109
15	*sec*-butylbenzene	2.816	2.788	2.730	2.754	1.100	1.174	1.269	1.341
16	*tert*-butylbenzene	2.903	2.981	2.920	2.957	0.900	0.753	0.796	0.867
17	fluorobenzene	3.086	3.086	3.086	3.086	0.590	0.590	0.590	0.590
18	chlorobenzene	2.018	2.120	2.369	2.155	6.901	5.467	5.039	3.082
19	bromobenzene	2.339	2.238	2.099	2.161	3.300	4.166	4.965	5.731
20	iodobenzene	2.535	2.535	2.555	2.535	2.101	2.101	2.101	2.004
21	1,2-dichlorobenzene	1.442	1.398	1.812	1.762	25.998	28.816	12.460	11.111
22	1,2-dibromobenzene	1.717	1.762	1.505	1.744	13.802	12.452	12.970	22.518
23	1,3-dichlorobenzene	2.477	2.471	2.418	2.312	2.400	2.432	3.508	2.749
24	1,3-dibromobenzene	1.717	1.723	2.050	1.779	13.802	13.619	11.974	6.420
25	1,2,4-trichlorobenzene	1.928	1.928	1.604	1.779	8.500	8.500	11.968	17.905

[a] The values are calculated for individual sets. Solvents 1−16 using eq 7, solvents 17−25 using eq 8. [b] Predicted using eq 13 for combined Set EF. [c] Predicted using eq 14 for combined Set EF. [d−f] Solubility in mg/mL obtained from a−c, respectively.
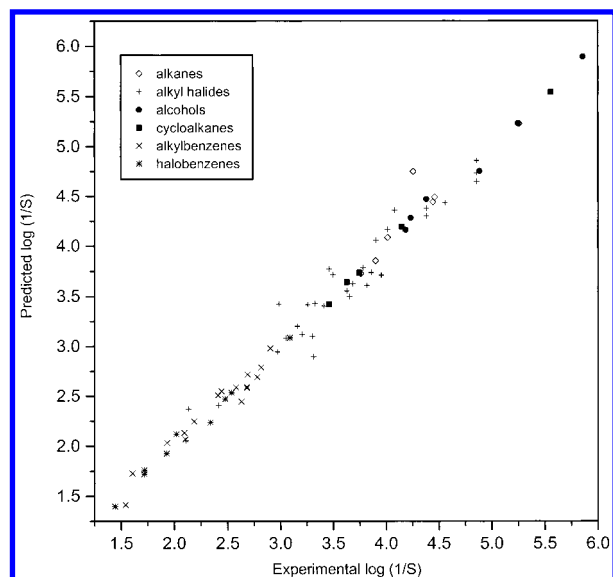
Figure 1. Correlation of experimental and predicted solubilities of C_{60}.

Figure 2. Comparison of experimental and predicted solubilities of C_{60} in mg/mL (log scale).

on the shape of the solvent molecule. Otherwise, the structure-solubility relation is dominant in cyclic solvents. Solubilities of C_{60} in aryl halides calculated using eq 8 are given in Table 7 (solvents 17−25, under the column "Pred[a]").

A comparison of the experimental solubilities with the solubilities predicted by the various regression equations is given in Figures 1 and 2.

Combined Sets. When Sets A and B were combined and the regression analysis procedures were repeated, no regression model with a standard error less than 0.2 was obtained. However, in several cases the predicted and experimental solubility matched very well. Wide deviation was observed

for dihaloalkanes. The results are given in Table 5 under the column "AB1" and "AB2", and the corresponding regression equations are eqs 9 and 10, respectively.

$$\log(1/S) = 12.428 - 35.325(\Phi) - 0.549({}^{4}\chi^{v}_{pc}) - 0.211({}^{3}\chi^{v}_{c}) - 1.377\,({}^{1}\chi) + 0.566({}^{3}\chi) + 0.792({}^{0}\chi^{v}) \quad (9)$$

$$n = 39, \ r = 0.952, \ r^2 = 0.906, \ SE = 0.241$$

where n, r, and SE are number of data, correlation coefficient, and standard error of the estimate, respectively.

$$\log(1/S) = 10.345 - 26.027(\Phi) - 0.478(^4\chi^v_{pc}) -$$
$$0.208(^3\chi^v_c) - 1.376\,(^1\chi) + 0.354(^3\chi) + 0.820(^0\chi^v) + IP$$
$$(10)$$

where IP = 0.316(F) $-$0.185(Cl) $-$ 0.467(Br) $-$ 0.652$-$
1.376(I), $n = 39$, $r = 0.960$, $r^2 = 0.922$, SE = 0.233.

For the combined Set ABC consisting of 45 solvents, polarizability parameter (Φ) and indicator parameter (IP) were sufficient to explain 91% of the data variability, but the standard error was 0.256. The regression model thus obtained is given in eq 11, and the values calculated were given in Table 5 under the column "ABC1"

$$\log(1/S) = 12.485 - 33.727(\Phi) + IP \qquad (11)$$

where IP = $-$ 0.07548(F) $-$ 0.0006196(Cl) + 0.469(Br) + 0.998(I) + 0.481(OH), $n = 45$, $r = 0.954$, $r^2 = 0.910$, SE = 0.256.

On including the χ values, there was no remarkable improvement in the standard error of the estimate. The regression equation including the χ values is

$$\log(1/S) = 12.225 - 32.211(\Phi) - 0.129(^4\chi^v_{pc}) -$$
$$0.0366(^2\chi) + IP \quad (12)$$

where IP = 0.03512(F) $-$ 0.007511(Cl) + 0.409(Br) + 0.880(I) + 0.432(OH), $n = 45$, $r = 0.958$, $r^2 = 0.917$, SE = 0.253.

The calculated solubility is given in Table 5 under the column "ABC2". Maximum deviations between the experimental and predicted values were observed in the case of dibromoalkanes.

When Sets E and F were combined, two regression models that have good predictability were obtained. As in the case of the individual Sets E and F, in the combined set also higher order χ terms correlated better with the solubility of C_{60}. The equation that includes only the higher order χ term is

$$\log(1/S) = 2.392 + 0.499(^6\chi^v_{pc}) + IP \qquad (13)$$

where IP = $-$0.637(o) $-$ 0.371(p) + 0.670(F) $-$ 0.09568-(Cl) $-$ 0.418(Br), $n = 25$, $r = 0.924$, $r^2 = 0.854$, SE = 0.217.

The other model includes higher order as well as lower order connectivity terms. The regression equation is

$$\log(1/S) = 2.155 - 2.392(^0\chi^v) + 3.461(^1\chi^v) +$$
$$1.549(^2\chi^v) - 0.4988(^6\chi^v_p) + IP \quad (14)$$

where IP = 0.03485(o) + 0.784(m) + 0.259(p) + 0.786(F) $-$ 0.07833(Cl) $-$ 0.000338(Br) + 0.423(I), $n = 25$, $r = 0.971$, $r^2 = 0.943$, SE = 0.207.

Solubility of C_{60} in the combined set of alkylbenzenes and aryl halides were calculated using eqs 13 and 14, and the solubility values are listed in Table 7 under the column "EF1" and "EF2", respectively.

VALIDATION STUDY

It had already been mentioned that one of the prime objective of this study was to develop a regression equation that could predict the solubility of C_{60} in complete range of organic solvents. Hence, it was decided to validate the regression equations of the combined models. However, this approach has not provided a regression equation which can be employed for calculating the solubility of fullerene in complete set of solvents. Hence the solvent lot was divided into aliphatic and aromatic sets. Again, within them, the total number of solvents were divided into training and testing sets, and the regression test was performed.

Data sets A, B, and C were combined to form the data set ABC i.e., aliphatic solvent set. The 45 solvents in this set were divided into training set (29 compounds i.e., alkanes 4, alkyl halides 21, and alcohols 4), and testing set (16 solvents). The regression equation obtained by the backward procedure with lowest standard error is given below with the statistics

$$\log(1/S) = (2.519 \pm 1.438) - (33.777 \pm 5.850)\Phi -$$
$$(0.185 \pm 0.211)^4\chi^v_{pc} + (1.031 \pm 0.392)I +$$
$$(0.456 \pm 0.224)Br + (0.478 \pm 0.318)OH \quad (15)$$

$$n = 29, r = 0.965, s = 0.231, SE = 0.254$$

where n, r, s, and SE are number of data, correlation coefficient, standard deviation, and standard error of the estimate, respectively. Though eq 15 has five independent parameters, except for Φ and $^4\chi^v_{pc}$, one of the other parameters only would be used for a given solvent. Hence, the regression equation could be considered a three-parameter model for aliphatic solvents. Solubility of fullerene, calculated using eq 15 for the training set are given in Table 8. The equation was validated by calculating the solubility of fullerene in the 16 solvents of the test set. Calculated solubility and experimental solubility as log(1/S) and mg/mL were given in Table 8 (Pred[a] gives log(1/S) and Pred[b] gives solubility in mg/mL). Good aggrement was observed between the predicted and experimental solubilities. However solvents such as 1,2-dibromopropane and 1,2-dicholoroethane showed significant deviation.

An attempt was also made to evolve a generalized equation for the aromatic solvents by combining data sets E and F. The combined data set was divided into a training set (17 solvents) and a testing set (eight solvents). While good agreement between experimental and predicted solubility was observed for ethylbenzene, sec-butylbenzene, and 1,3-dichlorobenzene, significant deviation was observed in the testing solvent set for four solvents, in particular chloro and bromobenzenes. The poor correlation is attributed to the limited set of data (only 25 solvents). To obtain a better picture, a greater number of solvents must be included in the training set, for which experimental data are not available as of now.

CONCLUSIONS

It is gratifying to note that the regression equations proposed for an individual class of solvents have reasonably good predictability to calculate the solubility of C_{60}. Though the regression models suggested for the combined data sets have a higher standard error, it is due to a few outliers. Though polarizability parameter and topological indices explained the major amount of data variations, inclusion of

1074 *J. Chem. Inf. Comput. Sci., Vol. 41, No. 4, 2001*

SIVARAMAN ET AL.

Table 8. Experimental and Predicted Solubility of C_{60} in Aliphatic Solvents (Set ABC)

no.	name	log(1/S) exptl	log(1/S) Pred[a]	solubility in mg/mL exptl	solubility in mg/mL Pred[b]
		Training Set			
1	pentane	5.255	5.101	0.004	0.006
2	octane	4.459	4.477	0.025	0.024
3	isooctane	4.442	4.214	0.026	0.044
4	dodecane	3.898	3.936	0.091	0.083
5	iodomethane	2.971	3.083	0.770	0.595
6	bromoethane	4.012	4.365	0.070	0.031
7	1-chloropropane	4.556	4.547	0.020	0.020
8	1-iodopropane	3.627	3.549	0.170	0.204
9	2-bromopropane	4.380	4.338	0.030	0.033
10	2-iodopropane	3.816	3.647	0.110	0.162
11	1-bromo-2-methylpropane	3.903	4.014	0.090	0.070
12	1-iodo-2-methylpropane	3.326	3.515	0.340	0.220
13	2-bromo-2-methylpropane	4.079	4.287	0.060	0.037
14	dibromomethane	3.301	2.818	0.360	1.095
15	bromochloromethane	2.982	3.328	0.750	0.338
16	1,2-dibromoethane	3.158	2.858	0.500	0.998
17	1,2-dichloropropane	3.857	3.549	0.100	0.203
18	1,3-dibromopropane	3.255	3.122	0.400	0.544
19	1,3-diiodopropane	2.415	2.388	2.770	2.948
20	dichloromethane	3.459	3.899	0.250	0.091
21	bromoform	2.106	2.392	5.641	2.919
22	1,1,1-trichloroethane	3.681	3.676	0.150	0.152
23	1,2,3-trichloropropane	3.051	2.749	0.640	1.284
24	1,1,2-trichlorotrifluoroethane	4.857	4.796	0.010	0.012
25	tetrachloroethane	2.133	2.336	5.301	3.325
26	ethanol	5.857	5.542	0.001	0.002
27	1-butanol	4.884	4.836	0.009	0.011
28	1-pentanol	4.380	4.640	0.030	0.016
29	1-octanol	4.185	4.289	0.047	0.037
		Testing Set			
1	hexane	4.255	4.696	0.040	0.014
2	decane	4.012	4.132	0.070	0.053
3	tetradecane	3.757	3.828	0.126	0.107
4	iodoethane	3.410	3.397	0.280	0.289
5	1-bromopropane	4.158	4.186	0.050	0.047
6	2-chloropropane	4.857	4.733	0.010	0.013
7	1-choloro-2-methylpropane	4.380	4.290	0.030	0.037
8	2-chloro-2-methypropane	4.857	4.608	0.010	0.018
9	2-iodo-2-methylpropane	3.496	3.799	0.230	0.114
10	1,2-dichloroethane	3.954	3.551	0.080	0.202
11	1,2-dibromopropane	3.313	2.889	0.350	0.930
12	1,3-dichloropropane	3.778	3.473	0.120	0.242
13	chloroform	3.653	3.510	0.160	0.222
14	carbontetrachloride	3.204	3.267	0.450	0.389
15	1-propanol	5.244	5.106	0.004	0.006
16	1-hexanol	4.234	4.485	0.042	0.024

an indicator parameter was essential in several data sets. This indicates that the variation in neither polarizability parameter nor structural parameters encodes the changes in solvent−solute interaction due to changes in substituents, their combination, and substitution pattern in the carbon skeleton. We continue our efforts in the evolution of a unified model that includes all solvents.

ACKNOWLEDGMENT

REFERENCES AND NOTES

(1) Sivaraman, N.; Dhamodaran, R.; Kaliappan, I.; Srinivasan, T. G.; Vasudeva Rao, P. R.; Mathews, C. K. Solubility of C_{60} in Organic Solvents. *J. Org. Chem*. **1992**, *57*, 6077−6079.

(2) Sivaraman, N.; Dhamodaran, R.; Kaliappan, I.; Srinivasan, T. G.; Vasudeva Rao, P. R.; Mathews, C. K. Solubility of C_{70} in Organic Solvents. *Full. Sci. Tech*. **1994**, *2(3)*, 233−246.

(3) Sivaraman, N.; Dhamodaran, R.; Kaliappan, I.; Srinivasan, T. G.; Vasudeva Rao, P. R.; Mathews, C. K. *Recent Advances in The Chemistry and Physics of Fullerenes and Related Materials*; Kadish, K. M., Ruoff, R. S., Eds.: U.S.A., 1994.

(4) Ruoff, R. S.; Tse, D. S.; Malhotra, R.; Lorents, D. C. Solubility of C_{60} in a variety of Solvents. *J. Phys. Chem*. **1993**, *97*, 3379−3383.

(5) Scrivens, W. A.; Tour, J. M. Potent Solvents for C_{60} and their Utility for the Rapid Acquisition of C^{13} NMR data for Fullerenes. *J. Chem. Soc. Chem. Commun*. **1993**, 1207−1209.

(6) Scrivens, W. A.; Cassell, A. M.; Kinsey, K. E.; Tour, J. M. In *Recent Advances in the Chemistry and Physics of Fullerenes and Related Materials*; Kadish, K. M., Ruoff, R. S., Eds.; U.S.A., 1994.

(7) Beck, M. T.; Mandi, G.; Keki, S. In *Fullerenes Vol. 2*; Ruoff, R. S., Kadish, K. S., Eds.; The Electrochem. Soc.: Pennington, NJ, 1995.

(8) Beck, M. T.; Mandi, G. In *Fullerenes Vol. 3*; Ruoff, R. S., Kadish, K. S., Eds.; The Electrochem. Soc.: Pennington, NJ, 1996.

(9) Beck, M. T.; Mandi, G. Solubility of C_{60}. *Full. Sci Tech*. **1997**, *5(2)*, 291−310.

(10) Kamlet, M. J.; Doherty, R. M.; Abboud, J. M.; Abraham, M. H.; Taft, R. W. Linear Solvation Energy Relatioships: 36. Molecular Properties Governing Solubilities of Organic Nonelectrolytes in Water. *J. Pharm. Sci*. **1986**, *75*, 338−349.

(11) Reinhard, M.; Drefahl, A. *Handbook for Estimating Physicochemical Properties of Organic Compounds*; John Wiley: New York, 1999; pp 118−139.

(12) 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.

(13) (a) Kier, L. B.; Hall, L. H. Molecular Connectivity in Chemistry and Drug Research, Academic Press: New York, 1976. (b) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis*; Research Studies Press: Chichester, England, 1986. (c) Basak, S. C.; Gute, B. D. Characterization of Molecular Structures Using Topological Indices. *SAR QSAR Environ. Res*. **1977**, *7*, 1−21. (d) *Topological Indices and Related Descriptors in QSAR and QSPR*; Devillers, J., Balaban, A. T., Eds.; Gordon and Breach: The Netherlands, 1999. (e) Basak, S. C.; Gute, B. D.; Grunwald, G. D. Use of Topostructural, Topochemical, and Geometric Prarmeters in the Prediction of Vapor Pressure: A Hierarchical QSAR Approach. *J. Chem. Inf. Comput. Sci*. **1997**, *37*, 651−655.

(14) Natarajan, R.; Anbazhagan, T. M.; Murali, T. C. Topological Calculator. To be submitted for publication.

(15) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976; pp 242−245.

(16) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976; p 249.

(17) (a) Nirmalakhandan, N. N.; Speece, E. R. Prediction of Aqueous Solubility of Organic Chemical Based on Molecular Structure. *Environ. Sci. Technol*. **1988**, *22*, 328−338. (b) Nirmalakhandan, N. N.; Speece, E. R. Prediction of Aqueous Solubility of Organic Chemical Based on Molecular Structure. 2. Application to PNAs, PCBs, PCDDs, etc. *Environ. Sci. Technol*. **1989**, *23*, 708−713.