New alternatives for estimating the octanol/water partition coefficient and water solubility for volatile organic compounds using GLC data (Kovàts retention indices)

Felicia Spafiu,[a] Alice Mischie,[a] Petre Ionita,[a] Adrian Beteringhe,[a] Titus Constantinescu,[a] and Alexandru T. Balaban[b,*]

[a]Institute of Physical Chemistry "Ilie Murgulescu" of the Romanian Academy, Splaiul Independentei 202, 060021 Bucharest, Romania
[b]Texas A&M University at Galveston, 5007 Avenue U, Galveston, TX 77551, USA
E-mail: balabana@tamug.edu, titelconstantinescu@yahoo.com

Abstract

New possibilities for estimating octanol/water partition coefficients ($\log P$) and the water solubility (S_w) were investigated using Kovàts retention indices (I) obtained from GLC retention data for 132 volatile organic compounds belonging to 7 different chemical classes (hydrocarbons, alcohols, aldehydes, ketones, carboxylic acids, esters and halogen compounds). Application of the multilinear regression method led to six equations, all involving index I, as follows: (*i*) direct correlation $\log P$ *vs. I* (eq.1); (*ii*) $\log P$ *vs. I* , molar refractivity, and surface tension (eq. 2); (*iii*) $\log P$ *vs. I* and structural characteristics (eqs. 3, 4 and 6); (*iv*) $\log P$ *vs.* the I/S_w ratio (eq. 5). Excepting eq. 1 (which showed relatively weak correlations), eqs. $2 - 6$ can provide reliable values for $\log P$ and for $\log S_w$ as proved by the significant statistical parameters. The general models presented through eqs. $2 - 6$ may also be applied in estimation of other biological and/or ecological important properties, which are linearly dependent on the $\log P$ or $\log S_w$ values. By generalization, a new calculation method is suggested (eq. 7), in order to allow the estimation of $\log P$ or $\log S_w$ in terms of the number of bonds and the Kovàts retention index.

Keywords: Log*P*, water solubility, Kovàts index, physical properties, QSPR

Introduction

The estimation of hydrophobic/hydrophilic properties of chemical compounds is relevant for many fields including medicine, pharmacology, foods, fragrances, chemical industry and environmental protection. These properties are measured by the octanol/water partition coefficient ($\log P$) and the water solubility under normal conditions (S_w)[1-10] both having current

applications in estimating various characteristics of interest for biology and environmental studies.[11-13]

In order to obtain experimental log$P$ values, liquid/liquid (l/l) extraction (octanol/water or, generally, organic solvent/water) may be applied. As an alternative to this time-consuming method, the Hansch model provides the log$P$ values from the experimental fragment and bond increments.[1,2] Besides this, several theoretical models for predicting log$P$ and logS_w values were proposed .[3-10,14-16] Another improvement of the classical l/l extraction is offered by reversed-phase liquid chromatography (RP-HPLC)[17,18] and reversed-phase thin-layer chromatography (RP-TLC),[19-21] which are at present the most frequently used techniques for providing experimental values of the octanol/water partition coefficients. Such methods are based on the linear correlation between log$P$ values of the corresponding compounds and the capacity factors (for RP-HPLC) or RM$_0$ values (in the case of RP-TLC), when reliable correlations are obtained within homologous series. In the case of volatile compounds, gas liquid chromatography (GLC) may provide more suitable alternatives.

The retention index on a certain stationary phase[22] is the result of a gas-liquid partition process, so that it may contain information related to solvation. Such information can be extracted from GLC data, under the form of well-known solubility factors,[23-25] which are used for the calculation of various properties, including log$P$. Besides this indirect method, only few attempts of evaluating directly physico-chemical properties from GLC data have been reported,[26] such studies being restricted to polycyclic aromatic derivatives.

The aim of the present study is to investigate the possibility of estimating the hydrophobic/hydrophilic properties of organic compounds, namely log$P$ and logS_w, from Kovàts indices (I) for a set of 132 organic compounds from 7 different classes. In order to obtain relationships suitable for structurally diverse sets, the simple correlations log$P$ *vs. I* were improved by addition of other parameters, such as molar refractivity, surface tension, number of bonds, accessible polar surface, or water solubility.

Results and Discussion

The use of the index I for estimation of the log$P$ value and water solubility was carried out in the present study as a stepwise strategy, which allowed drawing up of some new linear relationships between the mentioned parameters and the index I. The predictability of log$P$ and logS_w through the performed models was tested by cross-validation, using the leave-one-out method.[27] The leave-one-out cross-validation coefficient (R^2_{CV}) was provided by the CODESSA program.[28] Additionally, the quality factor of the regressions (Q) was calculated as R/SE.[29,30]

The set of compounds and the corresponding literature-accessible values for log$P$,[1,31,32] S_w,[6,7,32] and I index[24] are presented in Table 1 (stationary phase polyphenyl ether).

Table 1. Experimental values for Kovàts index(I), *logP* and water solubility (S_w) for the studied set of compounds

	Compound	log$P$exp	I	S_w(mol/l)
Hydrocarbons				
1	ethane	1.81	200	2.007×10^{-3}
2	propane	2.36	300	1.418×10^{-3}
3	n-butane	2.89	400	1.055×10^{-3}
4	n-pentane	3.39	500	5.27×10^{-4}
5	n-hexane	4	600	1.1×10^{-4}
6	n-heptane	4.5	700	3.39×10^{-5}
7	n-octane	5.15	800	5.78×10^{-6}
8	n-nonane	5.65	900	1.72×10^{-6}
9	n-decane	6.25	1000	$3.65E \times 10^{-7}$
10	n-undecane	6.54	1100	2.81×10^{-8}
11	n-dodecane	6.8	1200	−
12	n-tridecane	7.5	1300	−
13	n-tetradecane	8	1400	−
14	2-methylpropane	2.76	358	8.4×10^{-4}
15	3-methylheptane	−	774	6.92×10^{-6}
16	2,4-dimethylpentane	−	608	5.49×10^{-5}
17	ethane	1.13	270	4.67×10^{-3}
18	propene	1.77	360	4.753×10^{-3}
19	1-butene	2.4	450	3.939×10^{-3}
20	1-pentene	2.8	552	2.11×10^{-3}
21	1-hexene	3.4	614	5.94×10^{-4}
22	1-heptene	3.99	723	1.85×10^{-4}
23	1-octene	4.57	819	3.65×10^{-5}
24	*E*-2-octene	4.44	849	−
25	2-ethylhexene	4.31	824	−
26	2-butyne	1.46	587	−
27	1-octyne	3.57	907	−
28	2-octyne	3.57	979	−
29	benzene	2.15	833	2.482×10^{-2}
30	toluene	2.73	938	5.709×10^{-3}
31	ethylbenzene	3.15	1035	1.592×10^{-3}

32	styrene	2.95	1094	2.976×10^{-3}
33	ethynylbenzene	2.4	1096	4.464×10^{-3}
34	*o*-xylene	3.12	1046	1.677×10^{-3}
35	*m*-xylene	3.2	1050	1.51×10^{-3}
36	*p*-xylene	3.15	1041	1.52×10^{-3}
37	mesitylene	3.42	1155	4.01×10^{-4}
38	α-pinene	4.83	1015	1.83×10^{-5}
39	cyclohexane	3.44	735	−
Alcohols				
40	methanol	-0.77	522	31.21
41	ethanol	-0.31	587	21.7
42	1-propanol	0.25	703	16.63
43	1-butanol	0.88	820	8.52×10^{-1}
44	1-pentanol	1.56	929	2.49×10^{-1}
45	1-hexanol	2.03	1032	5.7×10^{-2}
46	1-heptanol	2.72	1138	1.43×10^{-2}
47	1-octanol	3.07	1241	4.14×10^{-3}
48	1-nonanol	4.02	1350	9.7×10^{-4}
49	1-decanol	4.57	1439	2.34×10^{-4}
50	1-undecanol	4.72	1569	1.11×10^{-4}
51	1-dodecanol	5.13	1678	2.15×10^{-5}
52	2-methyl-1-butanol	1.16	886	3.36×10^{-1}
53	3-methyl-1-butanol	1.42	885	3.02×10^{-1}
54	allyl alcohol	0.17	715	17
55	2-hexen-1-ol	1.6	1046	−
56	2-butanol	0.61	737	2.44
57	2-hexanol	1.76	950	1.34×10^{-1}
58	2-propanol	0.05	613	16.63
59	3-hexanol	1.65	941	−
60	2-methyl-2-propanol	0.35	654	13.49
61	2-methyl-2-butanol	0.89	771	1.13
62	2-methyl-2-pentanol	1.39	868	−
63	3-methyl-3-pentanol	1.41	900	−
64	3-methyl-3-heptanol	−	1078	1.83×10^{-2}
65	α-terpineol	2.98	1426	4.603×10^{-3}
66	dimethylcyclohexanol	2.37	1173	−

67	cyclopentanol	0.78	996	1.5
Aldehydes				
68	acetaldehyde	-0.22	541	22.7
69	propionaldehyde	0.3	656	5.26
70	butyraldehyde	0.83	753	9.84×10^{-1}
71	hexanal	1.89	966	5.63×10^{-2}
72	heptanal	2.42	1066	1.09×10^{-2}
73	octanal	2.9	1167	4.36×10^{-3}
74	2-methyl-1-propanal	0.77	706	–
75	propenal	-0.01	659	3.78
76	*E*-2-butenal	0.52	870	2.58
77	*E*-2-hexenal	1.58	1077	–
78	benzaldehyde	1.48	1272	6.54×10^{-2}
Ketones				
79	acetone	-0.24	652	17
80	2-butanone	0.29	764	3.092
81	2-pentanone	0.91	857	4.99×10^{-1}
82	2-hexanone	1.38	966	1.74×10^{-1}
83	2-heptanone	1.98	1061	3.76×10^{-2}
84	2-octanone	2.37	1165	7.01×10^{-3}
85	2-nonanone	3.14	1264	2.60×10^{-3}
86	2-decanone	3.73	1366	4.91×10^{-4}
87	2-undecanone	4.09	1469	1.16×10^{-4}
88	2-dodecanone	4.55	1571	–
89	3-hexanone	1.45	945	1.46×10^{-1}
90	cyclopentanone	0.38	1050	–
91	cyclohexanone	0.81	1160	2.54×10^{-1}
92	cycloundecanone	3.66	1757	–
93	cyclododecanone	4.1	1875	–
94	acetophenone	1.58	1387	5.10×10^{-2}
95	Carvone	2.52	1566	8.72×10^{-3}
Carboxylic acids				
96	acetic acid	-0.17	734	16.65
97	propionic acid	0.33	886	13.49
98	butyric acid	0.79	985	6.80×10^{-1}
99	valeric acid	1.39	1108	2.34×10^{-1}

100	hexanoic acid	1.92	1233	8.83×10^{-2}
101	heptanoic acid	2.42	1358	2.16×10^{-2}
102	octanoic acid	3.05	1483	5.47×10^{-3}
103	nonanoic acid	3.42	1608	1.79×10^{-3}
104	decanoic acid	4.09	1733	3.59×10^{-4}
105	undecanoic acid	4.42	1858	2.80×10^{-4}
106	dodecanoic acid	4.6	1982	2.4×10^{-5}
107	2-methylpropionic acid	0.94	945	1.89
108	2-methylbutyric acid	1.18	1094	4.39×10^{-1}
109	3-methylbutyric acid	1.16	1042	3.97×10^{-1}
110	4-methylpentanoic acid	1.79	1148	−
Esters				
111	methyl acetate	0.18	666	3.28
112	ethyl acetate	0.73	751	9.07×10^{-1}
113	2-methylbutyl acetate	2.29	1058	−
114	propyl acetate	1.24	844	1.85×10^{-1}
115	butyl acetate	1.82	958	7.23×10^{-2}
116	3-methylbutyl acetate	2.25	1010	1.50×10^{-2}
117	propyl butyrate	2.15	1024	1.24×10^{-2}
118	methyl propionate	0.82	770	7.08×10^{-1}
119	propyl formate	0.83	762	2.49×10^{-1}
120	isobutyl isobutyrate	2.48	1012	−
121	isopentyl isovalerate	3.62	1217	6.93×10^{-3}
122	benzyl acetate	1.96	1378	2.06×10^{-2}
123	methyl benzoate	2.23	1442	1.54×10^{-2}
Halogen compounds				
124	1-fluorooctane	4.2	972	1.72×10^{-4}
125	1-bromoethane	1.61	665	8.25×10^{-2}
126	1-bromopentane	3.37	973	8.41×10^{-4}
127	2-bromooctane	4.89	1214	3.26×10^{-5}
128	iodomethane	1.51	698	9.72×10^{-2}
129	1-iodobutane	3.08	989	1.09×10^{-3}
130	2-iodobutane	3.05	946	1.14×10^{-3}
131	1-chlorohexane	3.66	968	2.95×10^{-3}
132	*o*-chlorotoluene	3.42	1273	−

1. Estimation of log *P* using the Kovàts retention index (*I*)

1.1. Simple correlation of the log P parameter with the Kovàts index(I)

As it may be observed from Figure 1, log*P* is linearly dependent on index *I* within various homologous series. Plotting log*P* *vs*. *I* also shows trends, depending on the functional group, unsaturation, and the presence of rings or aromatic structures.
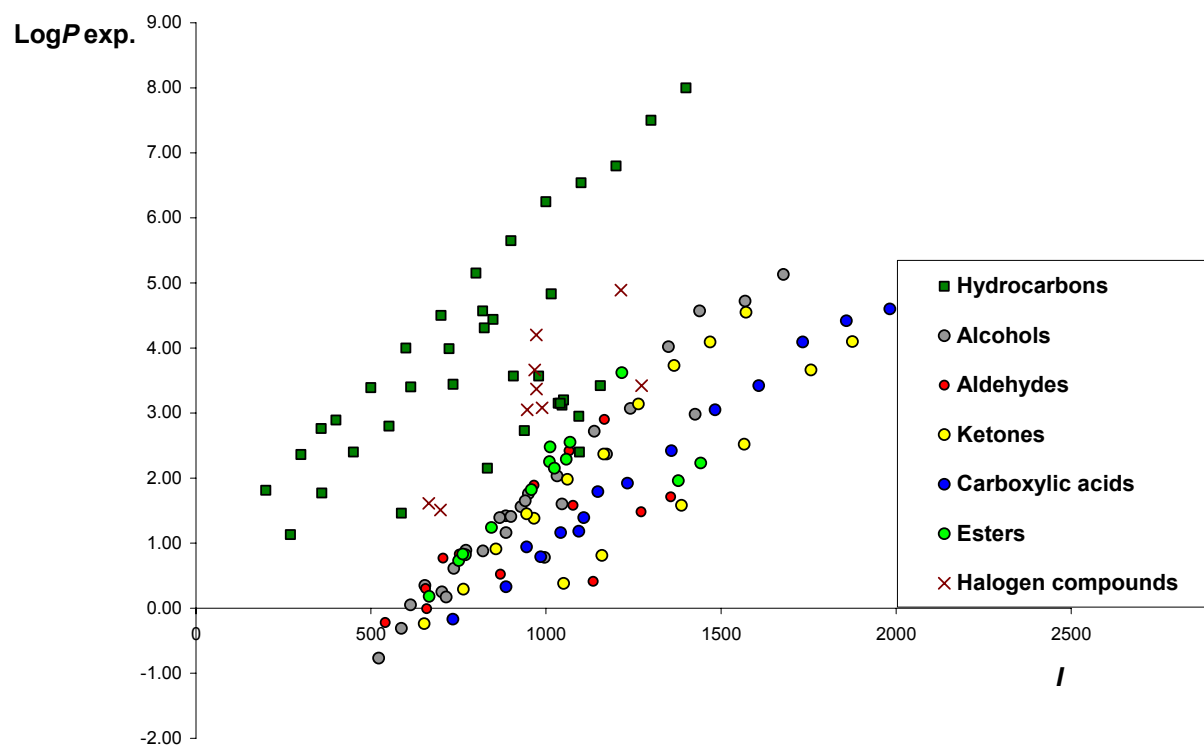

Figure 1. Plot of log*P* parameter (experimental, Table 1) *vs*. *I*.

In order to obtain linear relationships between log*P* and *I*, sets were formed according to functional groups, as shown in Table 1. Because the experimental *I* values presented in Table 1 are about two orders greater than log*P*, index *I* was replaced in the calculations with the ratio *I*/100, denoted as I^*. Linear regressions were performed according to a general equation (eq. 1), leading to relatively weak correlations ($R^2 = 0.431 - 0.990$). Statistics of these correlations are presented in Table 2.

$$\log P = a \times I^* + b \qquad \text{eq. 1}$$

where: $I^* = I$ (Kovàts index, Table 1)/ 100

Table 2. The coefficients *a, b* and statistical parameters [a] (R, F, SE, R^2_{CV}, Q) in the case of eq. 1, for each class of compounds (according to Table 1)

Eq.	Class [b]	N	*a*	*b*	R^2	F	SE	R^2_{CV}	Q
1a	Hydrocarbons	37	0.352	0.938	0.431	26.52	1.265	0.361	0.518
1b	Alcohols	27	0.493	-3.135	0.948	457.5	0.360	0.938	2.702
1c	Aldehydes	11	0.462	-1.913	0.703	21.35	0.569	0.508	1.472
1d	Ketones	17	0.369	-2.383	0.728	40.28	0.810	0.672	1.053
1e	Carboxylic acids	15	0.400	-3.034	0.990	1299	0.156	0.985	6.371
1f	Esters	13	0.292	-1.158	0.568	14.46	0.641	0.358	1.174
1g	Halogen comp.	9	0.446	-1.120	0.666	13.98	0.675	0.382	1.208

([a]) *R* = Pearson's correlation coefficient; *F* = Fisher test; *SE* = standard error; R^2_{CV} = cross validated coefficient; *Q* = quality factor.

There are significant differences between the correlation parameters corresponding to the various classes, but such differences may be due rather to the heterogeneous composition of the sets, with respect to the olefinic, cyclic or aromatic structures. The best correlation coefficients were observed in the case of alcohols (R^2=0.948) and carboxylic acids (R^2=0.990). Such pronounced linearity may be explained by the smaller proportion of unsaturated, cyclic or aromatic structures within these sets.

1.2. Dependence of the log P parameter on the Kovàts index(I) and physical properties (molar refractivity and surface tension)

Eq. 1 was improved, according to the general equation 2, by the addition of a second parameter, constructed from two physical properties (molar refractivity and surface tension) which affect the partition process between water and the organic solvent. The molar refractivity and surface tension were computed using the ACD ChemSketch 8.0 Freeware software.[33] Coefficients and statistical parameters of eqs. 2a – 2g, corresponding to each class of compounds, are presented in Table 3. These results show significant increases of the correlation coefficients (R^2 = 0.927 – 0.992). At the same time, a good correlation between log$P$ values calculated through eqs. 2a – 2g and the experimental log$P$ is displayed by Figure 2.

$$\log P = a \times I^* + b \times \log(MR/ST) + c \qquad \text{eq. 2}$$

where: MR = molar refractivity and ST = surface tension.

Table 3. Coefficients *a, b, c* and statistical parameters (R^2, *F*, *SE*, R^2_{CV}, *Q*)[a] in the case of eq. 2, for each class of compounds (according to Table 1)

Eq.	Class	N	*a*	*b*	*c*	R^2	*F*	*SE*	R^2_{CV}	*Q*
2a	Hydrocarbons	37	0.134	11.54	0.969	0.927	216.7	0.459	0.912	2.097
2b	Alcohols	27	0.402	2.218	-2.315	0.959	283.2	0.325	0.944	3.013
2c	Aldehydes	11	0.156	6.214	-0.183	0.938	61.48	0.274	0.890	3.534
2d	Ketones	17	0.160	7.145	-0.673	0.958	161.9	0.328	0.928	2.984
2e	Carboxylic acids	15	0.330	1.523	-2.113	0.992	839.4	0.137	0.986	7.270
2f	Esters	13	0.143	6.581	-0.153	0.970	165.5	0.175	0.922	5.627
2g	Halogen comp.	9	0.179	6.402	0.996	0.980	149.3	0.177	0.921	5.592

([a]) *R* = Pearson's correlation coefficient; *F* = Fisher test; *SE* = standard error; R^2_{CV} = cross validation coefficient; *Q* = quality factor.
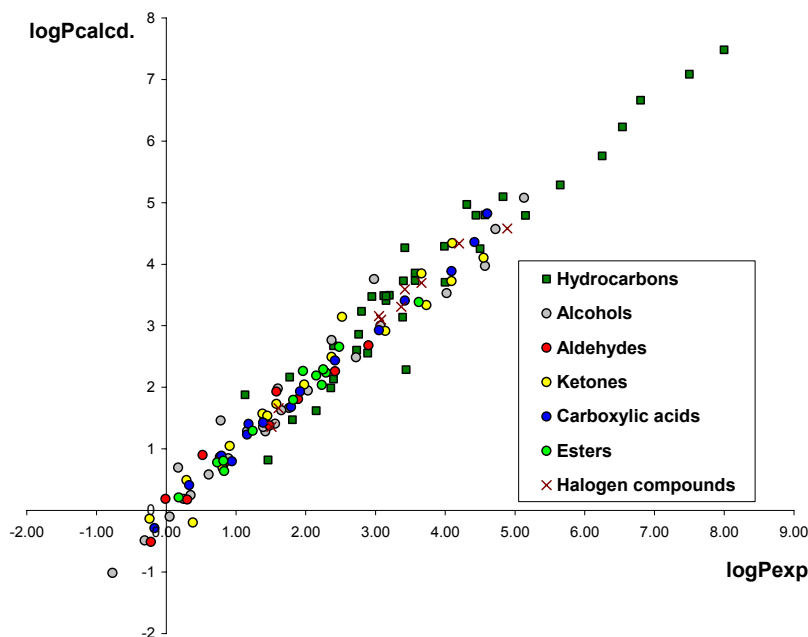

Figure 2. Plot of log*P* calcd. (eq.2, Table 3) *vs.* log*P* exp. (Table 1).

1.3. Dependence of the log P parameter on the Kovàts index involving structural effects and accessible polar surface

Another improvement of eq. 1 was achieved by a QSPR study. According to the observed influence of the molecular structure on the dependence log*P vs. I* (Figure 1), two parameters were added to eq.1: the number of bonds (nBt)[34] in order to characterize the effects of unsaturation, rings, or aromaticity, and the accessible polar surface, which may describe the availability of the functional group to hydration. Ascribing net atomic charges (CNDO) and

geometry optimization (MM$^+$ force field) were performed using the HyperChem program,[35] followed by computing of the accesssible surface of the heteroatoms.

As seen in Table 4 and Figure 3, the use of the number of bonds nBt (a constitutional descriptor) together with index I (eq. 3) led to a significant increase of statistical parameters, compared to eq.1.

$$\log P = a \times I^* + b \times nBt + c \qquad \text{eq. 3}$$

where nBt = number of bonds.

Table 4. The coefficients a, b and statistical parameters[a] (R, F, SE, R^2_{CV}, Q) in the case of eq. 3, for each class of compounds (according to Table 1)

Eq.	Class	N	a	b	c	R^2	F	SE	R^2_{CV}	Q
3a	Hydrocarbons	37	-0.118	0.213	0.517	0.990	1842	0.162	0.979	6.141
3b	Alcohols	27	0.183	0.119	-2.414	0.970	399.7	0.275	0.961	3.583
3c	Aldehydes	11	0.057	0.152	-1.418	0.988	334.4	0.120	0.976	8.284
3d	ketones	17	-0.031	0.187	-1.686	0.984	443.5	0.200	0.970	4.959
3e	Carboxylic acids	15	0.201	0.082	-2.170	0.993	889.5	0.133	0.988	7.493
3f	Esters	13	0.079	0.134	-1.586	0.975	200.1	0.160	0.915	6.171
3g	Halogen comp.	9	0.152	0.118	-0.084	0.982	171.3	0.165	0.923	6.005

([a]) R = Pearson's correlation coefficient; F = Fisher test; SE = standard error; R^2_{CV} = cross validation coefficient; Q = quality factor.
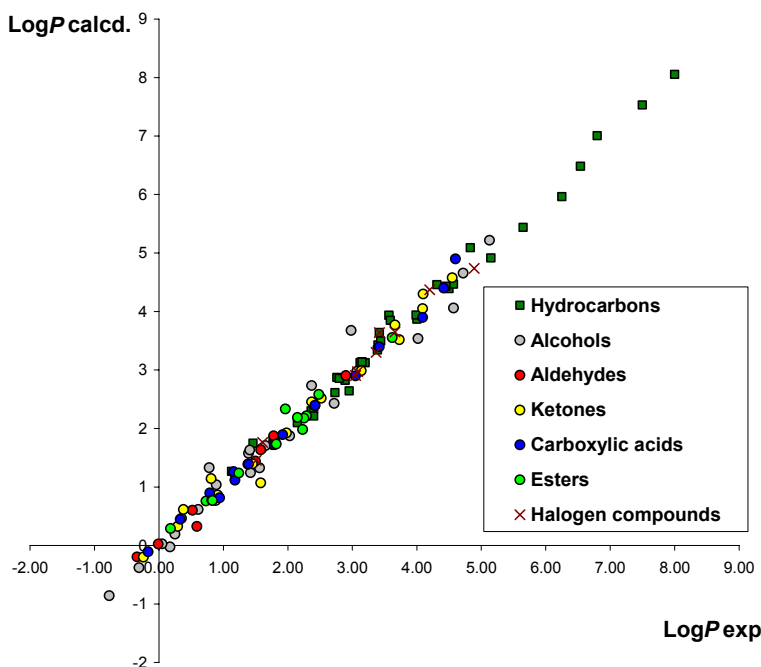

Figure 3. Plot of log$P$ calc (eq. 3, Table 4) *vs* .log$P$ exp. (Table 1).

Attempts were made for finding a unique relationship able to predict log$P$ of all the 129 compounds (for which experimental log$P$ was available, Table 1), independent of the type of the functional group. In this case, supplementary characteristics of the functional group are required for these compounds. The accessible polar surface was found to be the best from several tested descriptors. It was included in the QSPR model (eq. 4) as a substituent factor (SF), only in the case of heteroatom-containing structures. Satisfactory statistics could be achieved only by adding an indicator variable (V) to the QSPR model, so that V has value "1" for those structures which contain oxygen atoms, and "0" for the other ones.

$$\log P = 0.031 \times I^* + 0.185 \times nBt + 6.05 \times 10^{-3} \times SF + 0.353 \times V \qquad \text{eq.4}$$
$$N = 129; \ R^2 = 0.974; \ F = 1194.4, \ SE = 0.275, \ R^2_{CV} = 0.969, \ Q = 3.840$$

where: $I^* = I$ (Kovats index, Table 1)/100, nBt = number of bonds, SF = Sp = solvent accessible surface of the heteroatoms; for hydrocarbons, SF = 0; $V = 1$ for compounds containing oxygen atoms, and $V = 0$ for the other ones.

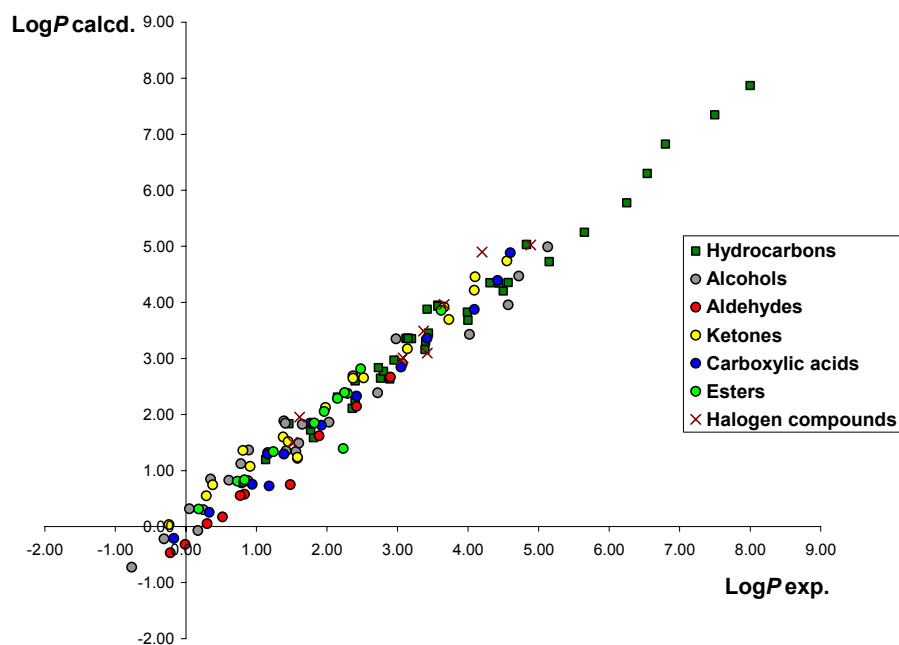In Figure 4 the log$P$calcd. values (calculated via eq. 4) are plotted against the experimental values log$P$exp.

Figure 4. Plot of log$P$calcd. (eq.4) *vs.* log$P$exp (Table 1)

Although R is slightly lower than in the case of eqs.3a – 3g, the statistical parameters of eq. 4 ($R^2 = 0.974$, $F = 1194.4$, $SE = 0.275$) may be considered relatively good, taking into account the number and the chemical diversity of the structures covered by this relationship.

When "1" value was assigned to variable V for both oxygen- and fluorine-substituted compounds, no improvement was observed ($R^2 = 0.971$, $F = 1065$, $SE = 0.291$). Taking into

account the hydrogen bond acceptor character of the fluorine atoms, an improvement of the statistics for equation 4 should be expected, if the variable V is related to hydrogen bonding. Such a result suggests that variable V is related rather to the Brønsted acid-base behavior of the oxygen-containing compounds.

1.4 Dependence of the logP parameter on index I and water solubility

As it is known,[1,2] log$P$ involves the partition of a chemical compound between an organic solvent (usually 1-octanol) and water. Starting from this basic notion, we attempted to correlate log$P$ values with the ratios I/S_w, where S_w represents the water solubility (eq. 5a, obtained using those compounds for which both the experimental log$P$ and S_w were available, Table 1). Such a strategy may offer an easily accessible method for measuring log$P$, being related to the real partition process (organic phase/water). A plot of log$P$ calculated (eq. 5a) and the experimental log$P$ is presented in Figure 5.
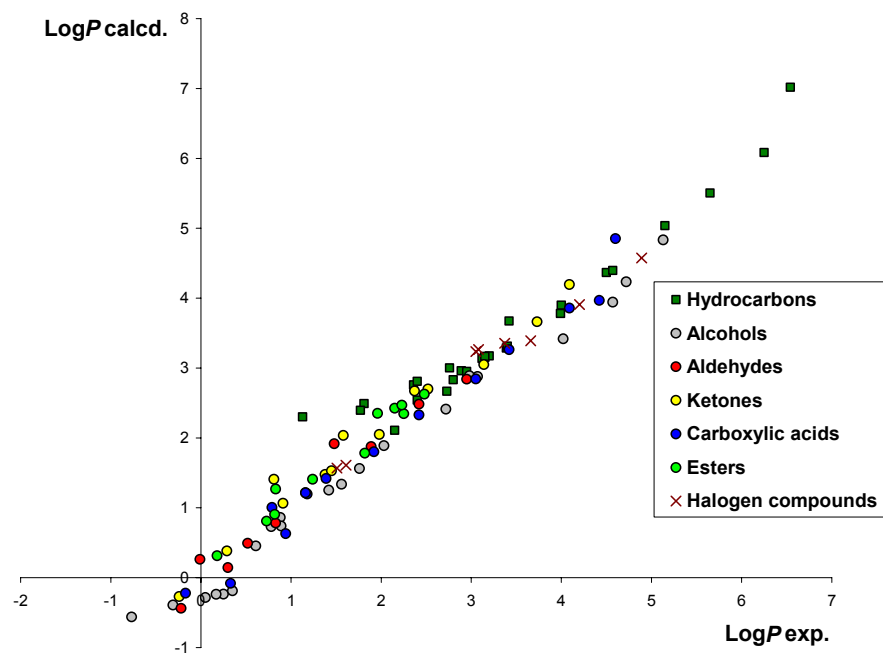
Figure 5. Plot of the log$P$ calcd. (eq.5a) *vs.* log P exp. (Table 1).

$$\log P = 0.809 \times \log(I^*/S_w) + 0.066 \qquad \text{eq. 5a}$$
$$N = 105;\ R^2 = 0.969;\ F = 3183,72;\ SE = 0.281;\ R^2_{CV} = 0.965;\ Q = 3.501$$

where S_w = experimental water solubility.

The resulted correlation equation (eq. 5a) was compared to Yalkowsky's equation applied to the same set of compounds (eq. 5b),[4] showing an improvement of statistical parameters, when the

Kovàts index is used together with water solubility. The improvement added by the retention index (which is related to the solubility in the organic phase) may depend on the stationary phase.

$$\log P = -0.823 \times \log Sw + 0.812 \qquad \text{eq.5b}$$
$$N = 105; R^2 = 0.950; F = 1981,84; SE = 0.353; R^2_{CV} = 0.949; Q = 2.762$$

2. Estimation of water solubility on the basis of Kovàts retention indices

The linear dependence between $\log P$ and S_w discussed above is due to the physico-chemical definition of the $\log P$ parameter, i. e. the organic phase/water partition. Based on such a relationship, $\log S_w$ may be expressed as a linear function of the same parameters that describe $\log P$. Thus, the general model represented by eq.3 (which gave the best results from all the tested models) was applied for estimation of $\log S_w$, resulting in eq. 6:

$$\log S_w = a \times I^* + b \times \text{nBt} + c \qquad \text{eq. 6}$$

where: $I^* = I$ (Kovàts index, Table 1)/100, nBt = number of bonds.

Table 5. Coefficients *a, b, c* and statistical parameters (*R, F, SE, R^2_{CV}, Q*) in the case of eq. 6, for each class of compounds (according to Table 1)

Eq.	Class	N	*a*	*b*	*c*	R^2	*F*	*SE*	R^2_{CV}	*Q*
6a	Hydrocarbons	30	0.217	−0.236	−0.93	0.931	184.4	0.372	0.905	2.593
6b	Alcohols	23	−0.222	−0.122	3.76	0.972	358.3	0.317	0.961	3.110
6c	Aldehydes	9	−0.255	−0.120	3.41	0.977	131.32	0.232	0.959	4.260
6d	Ketones	13	−0.045	−0.138	3.15	0.983	289.9	0.208	0.945	4.766
6e	Carboxylic acids	14	−0.267	−0.081	3.74	0.986	239.3	0.230	0.982	4.317
6f	Esters	11	−0.101	−0.144	2.45	0.952	79.62	0.220	0.898	4.435
6g	Halogen comp.	8	−0.408	0.068	2.05	0.987	201.47	0.156	0.954	6.368

(ᵃ) *R* = Pearson's correlation coefficient; *F* = Fisher test; *SE* = standard error; R^2_{CV} = cross validation coefficient; *Q* = quality factor.
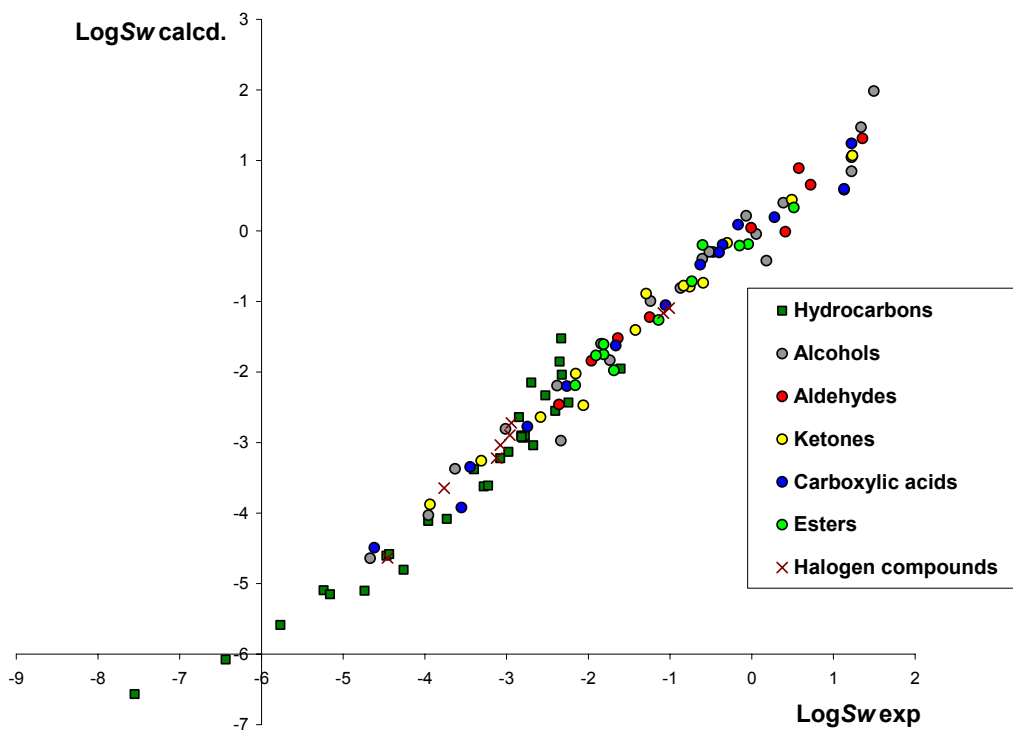
Figure 6. $\log S_w$ calcd. (eq. 6, Table 5), *vs*. $\log S_w$ exp.

The coefficients of eq. 6, corresponding to each class of compounds from table 1, and the statistical parameters are presented in Table 5. This result, together with the plot of $\log S_w$calcd *vs*. $\log S_w$exp. (Figure 6), augment the validity of such a model for estimating the water solubility of various compounds.

Because both parameters $\log P$ and $\log S_w$ can be expressed by the same QSPR model involving index I, a new general relationship, suitable for the estimation of $\log P$ or $\log S_w$, is proposed as eq. 7:

$$\log P \text{ or } \log S_w = a \times I^* + b \times nBt + c \qquad\qquad \text{eq.7}$$

where the definitons of I^*, and nBt are those specified for eq. 3 and eq. 6.

Values of the coefficients *a, b, c* have been already presented for each calculated property and for each class of compounds from Table 1.

The present study based on the correlation between $\log P$ and Kovàts index underlines the possibility of using GLC for the estimation of the octanol/water partition coefficients instead of liquid chromatography, in the case of volatile compounds. The linear relationship between these two parameters within homologous series could be extended to more general relationships, applicable to sets of increased structural diversity (eqs. 2 − 5). In the present work, several descriptors related to the physico-chemical properties of the solutes were tested, but many other descriptors may give significant results.[34,36]

The intercorrelation of parameters is presented in Table 6 for the most general models, covering large datasets, namely for eq. 4 (Table 6a) and eq. 5a (Table 6b).

Table 6. Intercorrelation of parameters for eqs. 4 (6a) and 5**a (6b)**

(6a)

	$\log P$	I^*	nBt	Sp	V
$\log P$	1				
I^*	0.458	1			
nBt	0.817	0.769	1		
Sp	-0.311	0.265	-0.147	1	
V	-0.523	0.337	0.030	0.436	1

(6b)

	$\log P$	$\log S_w$	$\log I^*$
$\log P$	1		
$\log S_w$	−0.975	1	
$\log I^*$	0.438	−0.329	1

Excepting the negative significant intercorrelation between $\log P$ and $\log S_w$, no other intercorrelation can be seen in Tables 6a and 6b. More pronounced intercorrelations of the descriptors may occur within particular sets from Table 1. Randić's orthogonalization procedure can be applied to the collinear parameters for each equation within each class of compounds. According to Randić,[37,38] orthogonalization of descriptors does not affect the statistical parameters (*R, F,* SE) which are the same for both orthogonal and nonorthogonal models. The main reasons for using orthogonal descriptors are the stability and significance of the equation's coefficients. Analysis of the parameters and of the coefficients of the presented equations, employing orthogonal models, on several stationary phases, and using larger data sets, will be the subject of a more detailed future study.

A comparison between the estimating abilities of the proposed models towards ClogP [1,31] and several known QSPR models (eg. AlogP,[14] MlogP [15] and ACDlabs [33]) can be made through the residuals between the experimental and the calculated $\log P$ and $\log S_w$ values, as presented in Table 7.

Table 7. Residuals of the experimental log*P* and log*S_w* values towards the corresponding values calculated through eq. 2 − 6, and through some already known QSPR models

		LogP							LogS_w	
	Compound	ClogP	AlogP[a]	MlogP[a]	Eq2	Eq.3	Eq.4	Eq.5	ACD	Eq.6
1	ethane	0.06	0.53	0.05	0.34	0.03	0.22	-0.68	0.69	-0.55
2	propane	0.08	0.62	0.08	0.37	0.06	0.25	-0.40	0.43	-0.21
3	n-butane	0.08	0.69	0.16	0.34	0.07	0.26	-0.07	0.12	0.15
4	n-pentane	0.05	0.74	0.25	0.25	0.04	0.23	0.11	-0.09	0.34
5	n-hexane	0.13	0.89	0.48	0.30	0.13	0.32	0.10	0.04	0.16
6	n-heptane	0.11	0.94	0.63	0.25	0.11	0.30	0.13	-0.04	0.14
7	n-octane	0.23	1.13	0.95	0.36	0.24	0.42	0.11	1.07	-0.14
8	n-nonane	0.20	1.17	1.13	0.36	0.21	0.40	0.15	-0.09	-0.18
9	n-decane	0.27	1.86	1.43	0.49	0.29	0.47	0.17	-0.13	-0.36
10	n-undecane	0.03	1.15	1.43	0.31	0.06	0.24	-0.48	0.17	-0.98
11	n-dodecane	-0.24	0.95	1.40	0.14	-0.21	-0.02	−	−	−
12	n-tridecane	-0.07	1.20	1.83	0.42	-0.03	0.15	−	−	−
13	n-tetradecane	-0.10	1.24	2.07	0.52	-0.05	0.13	−	−	−
14	2-methylpropane	0.08	0.77	0.03	-0.10	-0.11	0.11	-0.24	0.38	0.15
15	3-methylheptane	−	−	−	−	−	−	−	0.16	-0.01
16	2,4-dimethylpentane	−	−	−	−	−	−	−	0.09	0.54
17	ethene	-0.14	0.18	0.43	-0.75	-0.14	-0.06	-1.17	0.79	-0.81
18	propene	-0.02	0.46	0.55	-0.39	-0.03	0.05	-0.63	0.40	-0.29
19	1-butene	0.07	0.59	0.73	-0.27	0.06	0.15	-0.14	0.05	0.15
20	1-pentene	-0.05	0.54	0.72	-0.43	-0.06	0.03	-0.03	-0.18	0.36
21	1-hexene	0.02	0.68	0.94	-0.33	-0.02	0.09	0.09	-0.15	0.38
22	1-heptene	0.08	0.82	1.18	-0.30	0.05	0.16	0.21	-0.23	0.35
23	1-octene	0.13	0.94	1.42	-0.23	0.11	0.22	0.17	-0.15	0.15
24	*E*-2-octene	*0.01*	0.86	1.29	-0.36	0.02	0.10	−	−	−
25	2-ethylhexene	-0.02	0.69	1.16	-0.66	-0.15	-0.04	−	−	−
26	2-butyne	0.01	-0.41	-0.21	0.64	-0.29	-0.37	−	−	−
27	1-octyne	-0.03	-0.64	0.42	-0.16	-0.36	-0.38	−	−	−
28	2-octyne	-0.01	-0.13	0.42	-0.28	-0.26	-0.36	−	−	−
29	benzene	0.01	0.32	-0.11	0.53	0.05	-0.16	0.04	-0.35	0.35
30	toluene	0.09	0.41	0.12	0.13	0.12	-0.10	0.06	-0.17	0.19
31	ethylbenzene	-0.02	0.38	0.21	-0.26	0.01	-0.21	0.00	-0.16	0.13
32	styrene	0.09	0.57	0.10	-0.52	0.31	-0.02	0.00	0.09	-0.19
33	ethynylbenzene	-0.01	-0.56	-0.45	0.27	0.19	-0.20	-0.41	0.26	-0.50
34	*o*-xylene	0.03	0.32	0.18	-0.37	-0.01	-0.23	-0.02	-0.15	0.13
35	*m*-xylene	0.06	0.40	0.26	-0.29	0.08	-0.15	0.03	-0.10	0.08
36	*p*-xylene	0.01	0.35	0.21	-0.33	0.02	-0.21	-0.02	-0.10	0.10

37	mesitylene	-0.22	0.13	0.16	-0.85	-0.22	-0.46	-0.25	-0.06	-0.02
38	α-pinene	0.13	1.96	1.46	-0.27	-0.26	-0.20	0.12	0.50	0.36
39	cyclohexane	0.09	0.70	0.32	1.16	-0.06	-0.01	−	−	−
40	methanol	-0.01	-0.41	0.04	0.24	0.09	-0.05	-0.21	-0.62	-0.49
41	ethanol	-0.07	-0.30	-0.14	0.18	0.07	-0.09	0.08	-0.74	-0.14
42	1-propanol	-0.04	-0.27	-0.10	0.06	0.06	-0.05	0.49	-1.01	0.38
43	1-butanol	0.06	-0.09	0.08	0.05	0.11	0.06	0.02	-0.13	-0.29
44	1-pentanol	0.21	0.13	0.35	0.15	0.23	0.22	0.22	-0.04	-0.21
45	1-hexanol	0.15	0.15	0.44	0.08	0.15	0.17	0.14	0.14	-0.25
46	1-heptanol	0.31	0.38	0.78	0.23	0.29	0.33	0.31	0.26	-0.25
47	1-octanol	0.13	0.27	-0.12	0.07	0.09	0.16	0.19	0.29	-0.19
48	1-nonanol	0.55	0.77	0.52	0.49	0.48	0.59	0.60	0.37	-0.21
49	1-decanol	0.57	0.86	0.76	0.60	0.51	0.61	0.63	0.44	-0.26
50	1-undecanol	0.19	0.56	0.62	0.15	0.06	0.25	0.49	0.19	0.07
51	1-dodecanol	0.07	0.51	0.75	0.05	-0.09	0.14	0.30	0.28	-0.03
52	2-methyl-1-butanol	-0.06	-0.05	-0.13	-0.13	-0.09	-0.16	-0.06	0.00	-0.17
53	3-methyl-1-butanol	0.20	0.20	0.21	0.14	0.17	0.07	0.17	0.05	-0.22
54	allyl alcohol	0.16	-0.09	-0.03	-0.53	0.19	0.24	0.41	-0.92	0.17
55	2-hexen-1-ol	0.00	0.03	0.16	-0.38	-0.06	0.11	−	−	−
56	2-butanol	0.01	-0.28	-0.19	0.03	-0.01	-0.22	0.16	-0.42	-0.01
57	2-hexanol	0.10	-0.04	0.17	0.10	0.04	-0.09	0.20	-0.05	-0.06
58	2-propanol	-0.02	-0.32	-0.30	0.15	0.02	-0.27	0.33	-0.85	0.18
59	3-hexanol	-0.01	-0.22	0.06	0.02	-0.06	-0.17	−	−	−
60	2-methyl-2-propanol	-0.12	-0.22	-0.45	0.10	-0.11	-0.50	0.54	-1.01	0.54
61	2-methyl-2-butanol	-0.11	-0.21	-0.32	0.04	-0.15	-0.47	0.15	-0.37	0.10
62	2-methyl-2-pentanol	-0.14	-0.16	-0.20	0.03	-0.18	-0.50	−	−	−
63	3-methyl-3-pentanol	-0.12	-0.21	-0.18	-0.07	-0.22	-0.44	−	−	−
64	3-methyl-3-heptanol	−	−	−	−	−	−	−	-0.03	0.10
65	α-terpineol	0.35	0.57	0.62	-0.78	-0.70	-0.37	0.09	0.54	0.64
66	dimethylcyclohexanol	0.07	0.23	0.49	-0.40	-0.36	-0.32	−	−	−
67	cyclopentanol	0.07	-0.26	-0.04	-0.68	-0.55	-0.34	0.05	-0.19	0.60
68	acetaldehyde	-0.12	-0.04	0.10	0.29	-0.14	0.25	0.22	-0.57	0.04
69	propionaldehyde	0.29	-0.18	0.10	0.12	0.26	0.25	0.16	-0.55	0.07
70	butyraldehyde	0.05	-0.11	0.18	0.07	0.04	0.25	0.05	-0.41	-0.05
71	hexanal	-0.11	0.04	0.45	0.08	-0.09	0.27	0.02	-0.32	-0.03
72	heptanal	0.00	0.11	0.63	0.16	0.03	0.28	-0.06	-0.18	-0.12
73	octanal	-0.05	0.14	-0.14	0.22	0.00	0.23	0.11	-0.28	0.10
74	2-methyl-1-propanal	0.06	-0.18	0.12	-0.10	0.02	0.22	−	−	−
75	propenal	0.00	-0.53	-0.10	-0.20	-0.03	0.31	-0.27	-0.12	-0.31
76	E-2-butenal	-0.02	-0.40	-0.03	-0.38	-0.08	0.35	0.03	-0.70	0.42

77	*E*-2-hexenal	*-0.02*	-0.25	0.25	-0.35	-0.05	0.36	−	−	−
78	benzaldehyde	0.02	-0.11	-0.29	0.10	0.06	0.73	-0.44	-0.45	-0.12
79	acetone	-0.03	0.00	-0.44	-0.11	-0.04	-0.28	0.03	-0.97	0.16
80	2-butanone	-0.03	-0.13	-0.37	-0.20	-0.04	-0.26	-0.09	-0.69	0.05
81	2-pentanone	0.06	0.03	-0.15	-0.14	0.05	-0.17	-0.16	-0.36	-0.13
82	2-hexanone	0.00	0.04	-0.06	-0.19	-0.01	-0.22	-0.10	-0.36	0.03
83	2-heptanone	0.07	0.19	0.19	-0.07	0.06	-0.15	-0.07	-0.16	-0.02
84	2-octanone	-0.07	0.24	0.12	-0.12	-0.08	-0.28	-0.30	0.11	-0.13
85	2-nonanone	0.17	0.44	-0.22	0.22	0.15	-0.03	0.09	0.09	0.06
86	2-decanone	0.23	0.57	0.07	0.40	0.21	0.04	0.07	0.35	-0.05
87	2-undecanone	0.07	0.47	0.14	0.36	0.04	-0.13	-0.11	0.52	-0.06
88	2-dodecanone	0.00	0.48	0.32	0.45	-0.03	-0.19	−	−	−
89	3-hexanone	0.07	-0.10	0.01	-0.09	0.05	-0.07	-0.08	-0.29	-0.06
90	cyclopentanone	0.07	-0.34	-0.29	0.58	-0.23	-0.36	−	−	−
91	cyclohexanone	-0.05	-0.36	-0.24	0.12	-0.33	-0.55	-0.60	-0.08	0.14
92	cycloundecanone	0.00	0.21	1.01	-0.19	-0.11	-0.26	−	−	−
93	cyclododecanone	-0.12	0.19	1.17	-0.24	-0.20	-0.35	−	−	−
94	acetophenone	0.00	0.01	-0.52	-0.15	0.51	0.34	0.05	-0.86	-0.41
95	carvone	0.32	0.16	0.37	-0.63	0.00	-0.13	0.41	0.17	0.41
96	acetic acid	0.02	0.06	0.22	0.12	-0.06	0.04	-0.21	0.00	-0.02
97	propionic acid	0.00	-0.11	0.20	-0.08	-0.12	0.07	-0.03	-0.52	0.53
98	butyric acid	-0.07	-0.10	0.20	-0.10	-0.11	0.01	0.12	0.17	-0.26
99	valeric acid	0.00	0.04	0.39	-0.04	0.00	0.10	0.09	0.06	-0.15
100	hexanoic acid	0.00	0.11	0.55	-0.01	0.03	0.11	0.21	-0.06	0.00
101	heptanoic acid	-0.03	0.16	0.69	-0.02	0.02	0.09	0.16	0.03	-0.04
102	octanoic acid	0.07	0.33	0.08	0.12	0.15	0.21	0.23	0.15	-0.06
103	nonanoic acid	-0.09	0.25	0.13	0.01	0.02	0.06	0.45	0.18	0.03
104	decanoic acid	0.05	0.46	0.50	0.20	0.19	0.22	-0.25	0.45	-0.10
105	undecanoic acid	-0.15	0.33	0.54	0.06	0.02	0.03	0.31	0.18	0.37
106	dodecanoic acid	-0.50	0.06	0.44	-0.22	-0.30	-0.28	-0.02	0.87	-0.13
107	2-methylpropionic acid	0.30	0.35	0.04	0.14	0.12	0.18	-0.05	-0.19	0.08
108	2-methylbutyric acid	0.00	-0.18	0.18	-0.23	0.06	0.46	0.05	-0.12	-0.16
109	3-methylbutyric acid	-0.10	0.01	0.16	-0.07	-0.10	-0.13	0.41	-0.08	-0.10
110	4-methylpentanoic acid	0.00	0.19	0.42	0.11	0.07	-0.05	−	−	−
111	methyl acetate	0.00	0.16	0.05	-0.03	-0.11	-0.13	-0.13	-0.20	0.18
112	ethyl acetate	0.02	0.36	0.14	-0.05	-0.03	-0.08	-0.08	-0.20	0.14
113	2-methylbutyl acetate	-0.01	0.48	0.56	0.05	0.07	-0.09	−	−	−
114	propyl acetate	0.00	0.35	0.24	-0.06	0.00	-0.10	-0.17	-0.06	-0.02
115	butyl acetate	0.05	0.47	0.45	0.02	0.09	-0.03	0.04	-0.16	0.12
116	3-methylbutyl acetate	-0.05	0.65	0.52	-0.04	0.07	-0.14	-0.10	0.11	-0.06

117	propyl butyrate	1.44	0.13	0.42	-0.04	-0.04	-0.14	-0.28	0.11	-0.14
118	methyl propionate	0.03	0.13	0.23	0.01	0.04	0.02	-0.08	-0.09	0.06
119	propyl formate	-1.42	-0.02	0.24	0.19	0.06	0.00	-0.44	0.29	-0.40
120	isobutyl isobutyrate	0.00	0.14	0.42	-0.18	-0.10	-0.33	-0.15	0.07	0.03
121	isopentyl isovalerate	0.00	0.64	0.94	0.23	0.07	-0.23	−	−	−
122	benzyl acetate	0.00	0.36	-0.12	-0.31	-0.37	-0.09	-0.39	-0.40	0.29
123	methyl benzoate	0.12	0.55	0.20	0.19	0.25	0.84	-0.24	-0.20	-0.21
124	1-florooctane	0.00	0.30	0.25	-0.14	-0.17	-0.70	0.29	-0.01	-0.12
125	1-bromoethane	0.00	0.54	0.59	-0.05	-0.15	-0.35	0.00	-0.03	0.08
126	1-bromopentane	0.18	-0.01	0.20	0.06	0.07	-0.12	0.02	0.12	-0.04
127	2-bromooctane	0.11	0.46	-0.16	0.31	0.15	-0.14	0.32	0.04	0.18
128	iodomethane	0.05	0.27	0.01	0.15	0.06	0.00	-0.06	0.01	0.08
129	1-iodobutane	0.03	0.59	0.59	-0.02	0.11	0.07	-0.19	0.16	-0.06
130	2-iodobutane	0.00	0.82	1.13	-0.11	0.15	0.08	-0.19	0.34	-0.22
131	1-chlorohexane	0.08	-0.01	0.15	-0.04	0.01	-0.30	0.27	-0.58	0.10
132	*o*-chlorotoluene	0.07	0.23	0.48	-0.17	-0.22	0.33	−	−	−
	Average absolute deviation	**0.11**	**0.39**	**0.41**	**0.23**	**0.12**	**0.21**	**0.16**	**0.24**	**0.16**

[a]Values for AlogP and MlogP were available by using the Dragon software.[39]

In order to compare the presented results with the data calculated by the mentioned softwares, an average absolute deviation for the considered set was calculated for each model.

According to Table 7, average absolute deviations are smaller for eqs. 2 − 5 than for AlogP and MolgP, and in the case of eq.3 they are almost equal with those for ClogP. For the calculation of $\log S_w$, eq. 6 shows better results compared to the ACDlabs software.

Conclusions

Studying the possibility of estimation the octanol-water partition coefficients and water solubility starting from GLC data (Kovàts retention indices) in the case of 132 volatile compounds, belonging to seven different chemical classes, six equations were formulated involving the Kovàts index and additional parameters, namely molar refractivity and surface tension (eq. 2), number of bonds (eq.3 and 6), accessible polar surface (eq.4) and the I/S_w ratio (eq. 5a). As a generalization of the performed study, a new model (eq. 7) was proposed in order to estimate either $\log P$ or $\log S_w$ in terms of the Kovàts index and number of bonds. The linear relationship between $\log P$ and the Kovàts index formulated through eqs. 2 – 6, with R^2 values between 0.927 and 0.993, proves that GLC may offer a facile alternative for the estimation of the octanol/water partition coefficient for volatile compounds. Gas-liquid chromatography (GLC) is a simple and widely used technique for which very small samples are needed, compared to other physico-chemical methods. Moreover, eq. 7 evidences the possibility of extending each of the presented

models in order to calculate other properties, which are linearly dependent on log*P,* such as bioconcentration factors[11,13] or soil adsorption coefficients.[12]

This new approach based on gas chromatography alleviates the recent concerns in estimating log*P* when using liquid chromatography.[17-21] Likewise in the case of HPLC and TLC, improvements in the accuracy of property determination by using GLC retention indices may be expected by refinements in selecting the appropriate stationary phase. Such aspects, together with refining the already obtained models, need further investigations.

Supplementary Information

This is available in a Table of the molecular descriptors used in the above correlations for all the volatile organic compounds that are discussed in the present communication.

References

1   Hansch, C.; Leo, A.: *Substituent Constants for Correlation Analysis in Chemistry and Bio*log*y.* Wiley: New York, 1979.
2   Leo, A.; Hansch, C.; Elkis, D. *Chem. Rev.* **1971**, *71***,** 525.
3   Chiou, C. T.; Schmedding, D. W.; Block, J. H. *J. Pharm. Sci.* **1981,** *70***,**1176.
4   Yalkowsky, S. H.; Valvani, S. C. *J. Pharm. Sci.* **1980,** *69***,** 912.
5   Chiou, C. T. *Environ. Sci. Technol.* **1982,** *16***,** 4.
6   Yalkowsky, S. H.; Valvani S.C.; Roseman, T. J. *J. Pharm. Sci.* **1983**, *72*, 866.
7   Delaney, J. S. *J. Chem. Inf. Comput. Sc.i* **2004**, *44*, 1000.
8   Banerjee, S. *Environ. Sci. Technol.* **1985**, *19*, 369.
9   Maikut, O. M.; Makitra, R. C.; Palchikova, E. T. *Russ. J. Gen. Chem.* **2007**, *77*, 2198.
10 Tewari, Y. B.; Miller, M. M.; Wasik, S. P.; Martire, D. E. *J. Chem. Eng. Data* **1982**, *27*, 451.
11 Papa, E.; Dearden, J. C.; Gramatica, P. *Chemosphere* **2007**, 67, 351.
12 Briggs, G. G. *J. Agric. Food. Chem.* **1981**, *29***,** 1050.
13 Gestel, V.; Oterman, K.; Canton, J. H. *Regulatory Toxico*log*y and Pharmaco*log*y* **1985**, *5***,**422.
14 Ghose, A. K.; Crippen, G. M. *J. Comp. Chem.* **1986**, *7***,** 565.
15 Moriguchi, I.; Hirono, S.; Liu, Q.; Nakagome, I.; Matsuchita, Y. *Chem. Pharm. Bull.* **1992**, *40***,** 127.
16 Jäntschi, A., Bolboaca, S. D. *Int. J. Quant. Chem.* **2007**, *107*, 1736.
17 Kaliszan, R.; Haber, P.; Baczec, T.; Danuta, S. *Pure App.l Chem.* **2001**, *73*, 1465.
18 Gianinis, C.; Tsantili-Kakoulidou, A. *J. Liq. Chromatogr. Rel. Technol.* **2008**, *31***,** 79.
19 Kossoy, A. D.; Risley, D. S.; Kleyle, R. M.; Nurok, D. *Anal. Chem.* **1992**, *64***,** 1345.

20   Cserhati, T.; Forgacs, E. *J. Chromatogr. A* **1994**, *660***,** 313.

21   Beteringhe, A.; Radutiu, A. C.; Bem, M.; Constantinescu, T.; Balaban, A. T. *Internet. Electron. J. Mo.l Des.* **2006**, *5***,** 237.

22   Kovàts, E. *Helv. Chim. Acta* **1958,** *41* 1915.

23   Abraham, M. H. *Chem. Soc. Rev.* **1993**, 73.

24   Patte, F.; Etcheto, M.; Laffort, P. *Anal. Chem .***1982**, *54***,** 2239.

25   Laffort, P.; Chauvin, F.; Dallos, A.; Callegari, P. *J. Chromatogr. A* **2005,** *1100***,** 90.

26   Wang, I. H.; Wong, P. K. *Chemosphere* **2003**, *50*, 499.

27   Cruciani G.; Baroni, M.; Clementi, S.; Constantino, G.; Riganelli, D.; Skagerberg, B. *J. Chemometrics* **1992**, *6*, 335.

28   Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *CODESSA: A Reference Manual (Version 2.0)*, Gainesville, Florida, 1994.

29   Beteringhe, A.; Balaban, A. T.; *ARKIVOC* **2004**, *(i)*, 163.

30   Beteringhe, A.; Filip, P.; Tarko, L. *ARKIVOC* **2005**, *(x)*, 45.

31   Bio Loom program & database[http://www.biobyte.com/bb/prod/bioloom.html].

32   Syrracuse Research Corporation: free demos & databases. [http://www. syrres.com/esc/free_demos.htm].

33   [http://www.acdlabs.com/download/chemsk.html].

34   Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors.* Wiley-VCH: New York, 2000.

35   HyperChem^TM release 6.01 for Windows.

36   Devillers, J.; Balaban, A. T. Eds. *Topological Indices and Related Descriptors in QSAR and QSPR*. Gordon and Breach; The Netherlands, 1999.

37   Randić, M. *J. Chem. Inf. Comput. Sci.* **1991,** *31*, 311.

38   Randić, M. *New J. Chem.* **2000**, *24***,** 165.

39   Dragon 5.5 evaluation version.[http://www.talete.mi.it/download/download.php].