Solvation Parameters. 2. A Simplified Molecular Topology to Generate Easily Optimized Values

Paul Laffort* and Pierre Héricourt

CNRS, Centre des Sciences du Goût UMR 5170, 15 rue Hugues Picardet, F.21000 Dijon, France

Received January 11, 2006

This paper describes a generalized method to establish the values of the five solvation parameters of solutes, which reflect, together with the five solvation parameters of solvents, the intermolecular forces involved in solutions and in some biological phenomena. The tool applied for this purpose is a simplified molecular topology (SMT), which principally takes into account, for each atom of a molecule, its nature, the nature of its bonds, and in some cases the nature of its first neighbors. The learning material used to weight the molecular features generated by SMT are two sets of experimentally determined solvation parameters, established in a previous work (Laffort et al. *J. Chromatogr., A* **2005**, *1100*, 90−107).

INTRODUCTION

It has been proposed to call the *solvation parameters* of solutes the product of the molar volume and the previously called partial solubility parameters of solutes (Abraham et al.[2]). (In previous studies of Laffort and co-workers,[3−5] the concept of solvation parameters was called *solubility factors.*) The solvation parameters are involved not only in basic solubility phenomena but also in some biological properties, as summarized in Laffort et al.[1] There is presently a general agreement of the scientists involved in this field to consider that five independent solvation parameters of solutes and five solvation parameters of solvents are needed and are sufficient for a complete characterization of the solutions. The parameters of solutes are of dispersion, orientation, polarizability−induction, acidity, and basicity. The dispersion parameter is proportional to the bulk of the molecule, whereas the four other parameters (also called *polar*) are independent of the bulk.

Laffort et al.[1] have recently proposed two sets of optimized values of solvation parameters of solutes, on the basis of experimental measurements. The first set comes from a study by Abraham[6] of 340 substances, with a complete set of values for 314 of them (the dispersion parameter is lacking for 26 substances). To improve the mutual independence of the parameters published by this author, the parameters of dispersion and orientation have been modified using an internal rearrangement of the original values, via two simple equations.

The second set of optimized solvation parameters proposed by Laffort et al.[1] has been established for 133 substances by derivation from retention indices in gas−liquid chromatography on five selected stationary phases. Because these phases have been synthesized in limited quantity and are, therefore, not commercially available, this method should be adapted and tested using more common phases, before being considered as a general method to determine new values of solvation parameters.

The aim of the present study is to try to provide an easy and general method based on the molecular structure, validated by the above-mentioned two experimental data sets, to extend the knowledge of solvation parameters to new substances.

The molecular structure can be tackled using several techniques, each one with different advantages and drawbacks. One of these techniques, molecular topology, consists of representing the logical architecture of the molecules, leaving aside the exact length of the interatomic bonds and their mutual angles. There are several types of molecular topologies. The most complete and precise is certainly the DARC system, proposed by Dubois[7] as a retrieval tool in a data bank of molecules' descriptions (for example, the CAS data bank). In brief, the principle of this method is to consider a given atom of the molecule, called a focus, as a starting point for generating progressively a chemical graph environment in concentric layers. The DARC system was further applied to structure−activity relationships in physical chemistry (Dubois et al.[8,9]). The difficulty with the DARC system in this later field is that the possible number of parameters (independent variables) for a wide category of substances is so high that the method must be limited to narrow families of compounds. For example, in the case of the prediction of molar volumes at 25 °C, topological learning using 300 aliphatic alkanes in the range of carbon atoms C6−C11 is only applicable to this type of compound (Dubois et al.[9]). In opposition of this very precise topology, a very simple technique is to assign an increment to each type of atom (carbon, hydrogen, oxygen, etc.), to count their number and to make a weighted sum. This method is sufficient, for example, to calculate the molar mass. Numerous topological indices and vectors have been described in the literature; a survey can be found in Dubois et al.[10] (Among the most fruitful tools in the structure−activity field is the generation of molecular fragments followed by pattern recognition applied to these fragments; a review can be found in Laffort.[11] This procedure can, however, always be suspected of being dependent on the sample of molecules under study.) The topological method applied in the present study, very

* Corresponding author e-mail: laffort@cesg.cnrs.fr.

10.1021/ci0600152 CCC: $33.50
Published on Web 06/08/2006

Table 1. Nature of Atomic and Structural Elements Used in the Molecular Topology of the Present Study[a]

structural element	bonds	topological features	subcategories
		BASIC	
carbon	≤4	*C0*, C1, C11, C111, C1111, C2, C12, C112, C22, C3, C13	
oxygen	≤2	O0, O11	
oxygen	≤2	O1	linked to C1, C11, C111, C1111, C112
oxygen	≤2	O2	linked to C12, C112, others
nitrogen trivalent	≤3	N0	
nitrogen trivalent	≤3	N1	linked to C112, others
nitrogen trivalent	≤3	N11	
nitrogen trivalent	≤3	N111, N12, N3	
nitrogen pentavalent	≤5	N122	
phosphor pentavalent	≤5	P122	
fluorine	=1	F1	linked to C1111, others
chlorine	=1	Cl1	linked to C111, C1111, others
bromine	=1	Br1	
iodine	=1	I1	
sulfur divalent	≤2	*S0*, S1, S11, S2	
sulfur hexavalent	≤6	S111111, *S1122*	
silicon	≤4	Si1111	
tin	≤4	Sn1111	
hydrogen	=1	H1 = sum (maximal bonds − explicit bonds)	
		ADDITIONAL	
POSPA		phenol ortho substituted with proton acceptor (N122, O11, Cl, Br, I)	
POSPN		phenol ortho substituted with N122 (a particular case of POSPA)	
NCO		N1 or N11 or N111 linked to (a carbon linked to O2)	

[a] See explanations in the text. Italicized elements are not involved in the present study but are often present in organic chemistry.

similar to that applied in previous studies by Laffort and Patte[3] and Chauvin and Laffort,[12] is an intermediate approach. It takes into account for each atom of a molecule its nature, the nature of its bonds, and in some cases the nature of its first neighbors (cf. Experimental and Data Processing section).

EXPERIMENTAL AND DATA PROCESSING

We have applied a stepwise regression analysis to the molecular features generated by the molecular topology hereafter described (independent variables), to predict the five solvation parameters of solutes experimentally determined and also hereafter specified (dependent variables).

1. Molecular Topology Applied in the Present Study. Each atom is provided with an index, comprising a series of digits. Their sum is at most equal to the molecule's valence. The values of the digits define the type of bonds (1 for a single, 2 for a double bond, etc.), but the bonds with hydrogen are not indicated. In addition, the immediate neighboring is in part considered for oxygen, nitrogen, and halogens. So, the possibilities for oxygen, for example, are the following: O0, O1, O11, and O2, with four subcategories for O1 (linked to C11, C111, C1111, and C112) and three subcategories for O2 (linked to C12, C112, and other cases).

In addition to the 43 atom characteristics and their immediate environment finally kept in the present study, we also consider three additional topological features to try to account for spatial proximities between proton donors and proton acceptors existing in some ortho derivatives of phenols and in some amides, which we call POSPA, POSPN, and NCO. The definitions of the 46 features are summarized in Table 1. It should be noted that a connectivity parameter due to Zamora,[13] called the "smallest set of smallest rings", applied in one of our previous studies (Chauvin and Laffort[12]), has not been selected here by the stepwise multilinear regressions used.

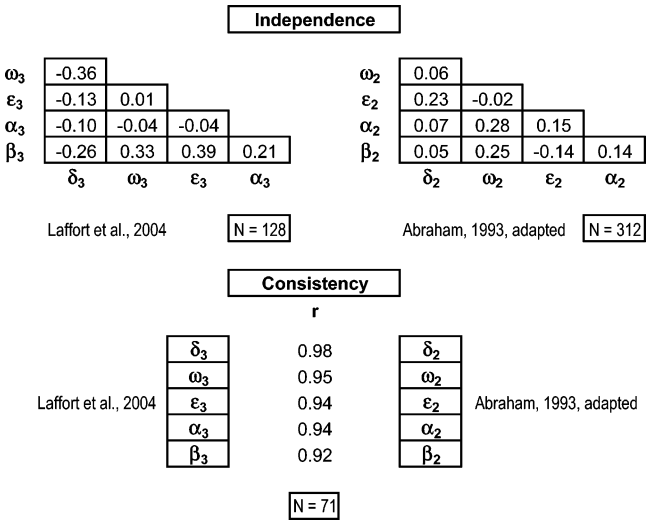

Figure 1. (top) Mutual independence of solvation parameters in the two data subsets (low values of correlation coefficients). (bottom) Good consistency for compounds present in both subsets (high values of correlation coefficients *r*).

In some ways, the strategy applied in the present study is a compromise between the assignments of fixed atomic increments and the consideration of all first and second neighbors of each structural element, as generally used with the DARC environment applied to structure−activity relationships.

Molecular topological features are generated by using an original algorithm called **S**implified **M**olecular **T**opology (SMT), based on the MarvinSketch program and other Java functionalities of ChemAxon Ltd.[14] In the first step, the molecules under study are drawn using the MarvinSketch

Table 2. Molecular Features and Their Corresponding Coefficients Involved in a Simplified Topology Defining the Five Solubility Parameters of Dispersion, Polarizability/Induction, Acidity, Orientation, and Basicity for the 369 Compounds Reported in Table A1[a]

DISPERSION

Features	Coefficients	Partial F ratios
Constant	-0.494	
C1+C11	0.494	12290
C tot-(C1+C11)	0.398	7249
O11	0.169	43
O tot-O11	0.491	800
N tot-N122	0.448	239
Cl1	0.615	767
Br1	0.911	766
I1	1.108	415
S tot	0.931	445
Si1111	0.610	74
Sn1111	1.295	44

POLARIZABILITY

Features	Coefficients	Partial F ratios
Constant	0.300	
C1	-0.150	515
C111	0.150	182
C1111	0.318	140
C12	0.055	182
C112	0.222	607
N111	0.250	35
F1	-0.237	561
Br1	0.152	73
I1	0.482	262
S tot	0.267	127
O2 x C112	-0.158	72

ACIDITY

Features	Coefficients	Partial F ratios
C3	0.293	59
O0	2.316	287
sqrt [O1 x (C1+C11)]	1.100	1031
O1 x C1111	0.655	67
O1 x C111	0.944	405
O1 x C112	1.817	5052
F x C1111	0.098	78
Cl x C111	0.133	22
N1 x C112	1.384	102
sqrt [(N0+N1+N11) except N1xC112]	0.401	125
POSPA	-0.899	221

ORIENTATION

Features	Coefficients	Partial F ratios
C12 + C112	0.023	45
O0+O1	0.213	78
O11	0.147	42
O2	0.474	771
N12	0.187	25
N3	1.022	566
F1	0.104	146
Cl1 except Cl1xC1111	0.188	83
Br1	0.159	30
NCO	0.687	59
POSPN	-0.747	20

BASICITY

Features	Coefficients	Partial F ratios
C112	0.052	39
N3	0.238	73
P1112	0.592	57
O0+O1 except O1xC112	0.359	386
O2x(C12+C112)	0.243	372
F1	-0.019	38
N1xC112	0.306	31
(N0+N1+N11+N111) except N1xC112	0.435	490
N12	-0.552	42
sqrt N12	1.096	109
sqrt O11	0.104	49

[a] The partial *F* ratios reflect the degree of involvement of each feature. The term "linked to" is marked x.

program. In the second step, the Java library of ChemAxon allows for scrolling through the molecular structures generated in the first step, to establish the needed structural parameters.

2. The Learning Experimental Data Set. The experimental data set permitting the establishment of the rules involving the 46 structural features reported in Table 1 results from the pooling of the two subsets already mentioned.

2.1. Values Derived from Retention Indices on Five Selected Stationary Phases in Gas−Liquid Chromatography by Laffort et al.[1] Concerning 133 Compounds. We discarded the compounds using retention chromatographic indices having a value < 500 (i.e., lower than that of pentane), for a better consistency. This first subset for 128 compounds is marked with the subscript 3.

2.2. Values Derived from the Data Published by Abraham. Out of the data published by Abraham[6] for 314 compounds, we discarded the data for teflurane and fluroxene, for which we did not find the molecular structure. The original data of Abraham[6] for the remaining 312 compounds were modified according to the following rules of transformation established by Laffort et al.,[1] the data being marked with a subscript 2:

$$\begin{aligned}
\text{DISPER} \quad & \delta_2 = \log L16 - 0.532\pi_2^{\text{H}} - 0.894R_2 - 0.115 \\
\text{ORIENT} \quad & \omega_2 = 1.523\pi_2^{\text{H}} - 0.538\Sigma\beta_2^{\text{H}} - 0.837R_2 \\
\text{POLARIZ} \quad & \epsilon_2 = R_2 \\
\text{ACID} \quad & \alpha_2 = 2.825\Sigma\alpha_2^{\text{H}} \\
\text{BASIC} \quad & \beta_2 = 0.728\Sigma\beta_2^{\text{H}}
\end{aligned} \tag{1}$$

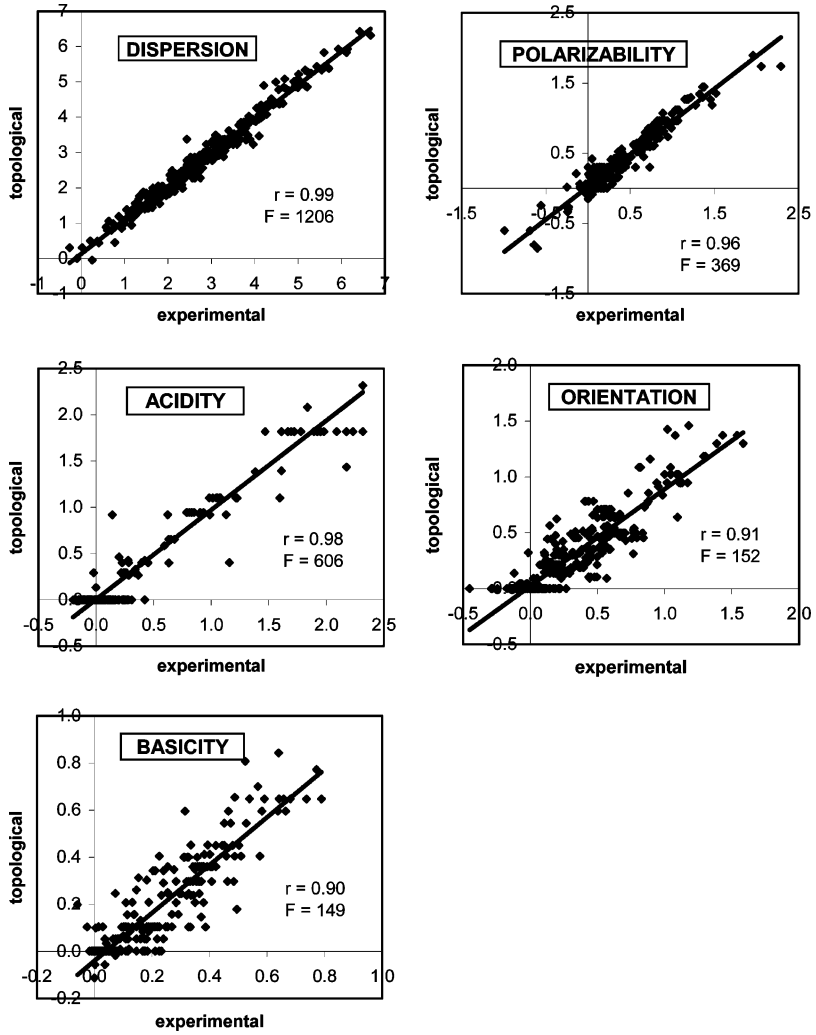

Figure 2. Correlograms of the solubility parameters topologically defined (Table 2) versus experimental results (Table A1). (The correlation for the acidity parameter repeated without the compounds with zero experimental values gives $N = 209$ and $r = 0.97$.)

in which log $L16$, π_2^H, R_2, $\Sigma\alpha_2^H$, and $\Sigma\beta_2^H$ are respectively the parameters of dispersion, orientation, polarizability, acidity, and basicity.

As shown by Laffort et al.,[1] the internal rearrangement of the original parameters providing the definitions of δ_2 and ω_2 allows good mutual independence of the solvation parameters, which was not the case with the original data. Multiplicative coefficients have been applied in the definitions of α_2 and β_2 in order to make the two subsets as comparable as possible.

These two subsets have a total of 369 defined compounds and an overlapping of 71 compounds. The good mutual independence of the parameters within the two data sets and the good consistency between both sets are shown in Figure 1. For the overlapping data, we applied the rule of preference of values from homogeneous chromatographic origin (i.e., from Laffort et al.[1]) in building the pooled set. The learning experimental data set for 369 compounds is listed in the table of Appendix A.

RESULTS

A stepwise multilinear regression analysis has been applied to the 369 compounds listed in the table of Appendix A, each of the five solubility parameters being successively the dependent variables and the 46 molecular features listed in Table 1 being the independent variables (taken alone or sometimes grouped, e.g., all types of oxygen features except O11). The first used rule has been to have not more than 11 final independent variables in each regression (which represents 3% of the total observations). The second rule has been that each partial F ratio associated with the independent variables finally kept has to be at least equal to 20. (The higher are the partial or global F ratios, the better is the fitting between the observed and predicted values. The F ratio depends of the correlation coefficient, the number of observations, and the number of independent variables.) The selected molecular features permitting the prediction of the five solubility parameters, their respective coefficients, and their partial F ratios are listed in Table 2. The corresponding correlograms are drawn in Figure 2. (The complete software based on the MarvinSketch program as well as the molecular features generated are available upon request: laffort@cesg.cnrs.fr or hericourt@cesg.cnrs.fr.)

It should be noted that the mutual independence of the solvation parameters for the 369 solutes studied here is almost identical when the experimental or the topological values are considered, as it can be observed in Figure 3.

It has also been mentioned that a popular parameter first proposed by Palm et al.[15] has been successfully applied in pharmacology, particularly to account for passive molecular

Table 3. Comparison of the Predicted Values of Acidity and Basicity Parameters in Two Recent Publications and in the Present Work vs Corresponding Experimental Data[a]

Jover et al. (2004)[19] (NN method II)	↔	experimental	↔	present work
	$N = 203$, $r = 095$, SEE = 0.17	α_2 from table A1	$r = 0.98$, SEE = 0.13	
	$N = 203$, $r = 0.93$, SEE = 0.06	β_2 from table A1	$r = 0.92$, SEE = 0.07	
Oliferenko et al. (2002)[26] (A* and B*)	↔	experimental	↔	present work
	$N = 37$, $r = 0.93$, SEE = 0.20	α_2 from table A1	$r = 0.97$, SEE = 0.13	
	$N = 43$, $r = 0.93$, SEE = 0.20	α_2 out of table A1	$r = 0.96$, SEE = 0.14	
	$N = 37$, $r = 0.91$, SEE = 0.07	β_2 from table A1	$r = 0.90$, SEE = 0.08	
	$N = 40$, $r = 0.97$, SEE = 0.07	β_2 out of table A1	$r = 0.95$, SEE = 0.09	

[a] *N*, *r*, and SEE stand respectively for the number of compounds, the correlation coefficient, and the standard error of estimate.

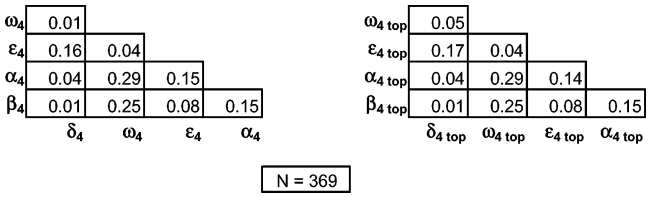

Figure 3. Mutual independence (i.e., low values of correlation coefficients) of the solvation parameters for the 369 solutes studied in the present work. The values are almost identical for experimental (on the left) and topological data (on the right).

transport through biological membranes (intestinal absorption, blood−brain barrier crossing, etc.). This parameter, defined as the sum of surfaces of polar areas, has been called the molecular potential surface area (PSA). Ertl et al.[16] suggested a new and fast protocol to calculate the PSA on the basis of the topological information only, which they termed TPSA.

We added TPSA values to the topological features of Table 1, to try to improve the fitting between experimental and topological solubility parameters. The TPSA values were never selected by the stepwise multilinear regression analysis. In other words, the linear equations of Table 2 could not be improved. It can be underlined that the difficulties previously mentioned, concerning some substituted phenols and some amides, are not overcome by using TPSA. By contrast, the TPSA values are very easily predicted from 11 of the *basic* molecular features of Table 1, for our 369 compounds ($r = 0.99$). (Ertl et al.[16] used 43 molecular fragments, but this is true for many more compounds.)

DISCUSSION

1. Various Definitions of Solvation Parameters. Over the past few years, an explosion of publications on solvation parameters, as well on their experimental determination, as in their evaluation from the molecular structure, has been observed, the former standing as a basis for the latter. All of these publications have a common concern: the solvation parameters as defined by Abraham.[6] However, the parameters defined in this way have not been validated individually but only by their aptitude to predict altogether other experimental properties (e.g., the partition coefficient of octanol−water, among many others), using linear multiple regressions (LMRA) or artificial neural networks (NN). In fact, Laffort et al.[1] have shown that the Abraham parameters are partially mutually correlated and that the parameters of dispersion and orientation need a new definition. In contrast, the parameters of induction−polarizability, acidity, and basicity according to Abraham[6] are validated. Laffort et al.[1] have also shown that the experimental determination of the solvation parameters can be 100% experimental, whereas the Abraham approach is only partially experimental.

2. Experimental Data Banks. The recovery of the parameters already published by Abraham and colleagues for a new definition of the dispersion parameter is limited to the solutes for which the partition coefficient air−hexadecane L16 is known. Therefore, the table in Appendix A currently appears to be the most complete on the basis of published data for the five parameters. The very interesting table published by Zissimos et al.,[17] for example, concerning 470 compounds, does not include values of the partition coefficient air−hexadecane L16 but only values of the molar volume Vx according to the definition of Abraham and Mc Gowan.[18] It should be interesting to further verify, using the tests applied by Laffort et al.[1], whether the product fn·Vx is equivalent or not to the parameter δ_2; this has not been done by these authors. (fn is the equation of the refractive index *n*: $\text{fn} = (n^2 - 1)/(n^2 + 1)$.) Some much more experimental data banks, such as those mentioned by Platts et al.[19] (1947−3692 solutes, according to the considered parameters), are only partially published. It is right that the table in Appendix A, as a basis of the validation of predictive models, does not include enough compounds with functional groups of more than one of the same type.

3. The Results Here Presented. The results summarized in Table 2 and Figure 2 are satisfactory for the parameters of dispersion and polarizability−induction: regular distribution of the data, high values of the correlation coefficients, and only topological features of *basic* type, as called in Table 1 (i.e., including only the nature of the atoms, their surrounding bonds except with hydrogen, and the nature of some first neighbors). By contrast, the results obtained for the three other solubility parameters (orientation, acidity, and basicity) are not so good: lower values of the correlation coefficients or an irregular distribution of the data. In addition, the so-called *basic* topological features of Table 1 are not sufficient to account for some internal hydrogen bonds, which modify one or another parameter of acidity, basicity, or orientation. In the collection of 369 compounds studied here, this phenomenon is clearly observed in some substituted phenols and in some amides. There is obviously a limitation, for these substances, of applying a too simple molecular topology. The *additional* topological features POSPA, POSPN, and NCO have been used to overcome this difficulty, but of course, this strategy has the same drawback as the strategy of molecular fragments mentioned previously: the procedure can be suspected to be dependent on the substances' sample under study.

Table A1. Learning Experimental Data Set of Solvation Parameters Values for 369 Solutes Used in the Present Study[a]

number	compounds	δ_4	ω_4	ϵ_4	α_4	β_4
		Alcohols				
1	water	−0.09	0.50	0.00	2.32	0.25
2	methanol	0.37	0.18	0.28	1.21	0.34
3	ethanol	0.93	0.18	0.25	1.05	0.35
4	1-propanol	1.48	0.18	0.24	1.05	0.35
5	2-propanol	1.27	0.07	0.21	0.93	0.41
6	1-butanol	***1.77***	***0.17***	***0.35***	***0.98***	***0.38***
7	2-butanol	***1.76***	***0.19***	***0.19***	***0.79***	***0.36***
8	1-pentanol	***2.32***	***0.16***	***0.33***	***1.01***	***0.39***
9	2-pentanol	***2.24***	***0.20***	***0.20***	***0.79***	***0.37***
10	2-methyl-2-butanol	***2.09***	***0.19***	***0.14***	***0.65***	***0.36***
11	1-hexanol	***2.88***	***0.16***	***0.29***	***1.06***	***0.39***
12	2-hexanol	***2.76***	***0.20***	***0.17***	***0.81***	***0.37***
13	2-methyl-2-pentanol	***2.52***	***0.21***	***0.15***	***0.63***	***0.36***
14	1-heptanol	***3.44***	***0.16***	***0.24***	***1.08***	***0.40***
15	2-heptanol	***3.27***	***0.20***	***0.17***	***0.83***	***0.38***
16	2-methyl-2-hexanol	***3.18***	***0.21***	***0.02***	***0.68***	***0.37***
17	2-phenylethanol	***3.45***	***0.38***	***0.72***	***1.22***	***0.40***
18	1-octanol	4.10	0.21	0.20	1.05	0.35
19	1-nonanol	4.61	0.22	0.19	1.05	0.35
20	1-decanol	5.12	0.22	0.19	1.05	0.35
21	1-undecanol	5.63	0.23	0.18	1.05	0.35
22	1-dodecanol	6.15	0.23	0.18	1.05	0.35
23	cyclopentanol	2.46	0.16	0.43	0.90	0.41
24	cyclohexanol	***2.85***	***0.08***	***0.51***	***0.85***	***0.41***
25	cycloheptanol	3.55	0.08	0.51	0.90	0.42
26	prop-2-en-1-ol	1.29	0.16	0.34	1.07	0.35
27	*trans*-but-2-en-1-ol	1.96	0.12	0.35	1.07	0.35
28	phenol	2.46	0.52	0.81	1.70	0.22
29	*o*-cresol	2.89	0.45	0.84	1.47	0.22
30	*m*-cresol	2.99	0.47	0.82	1.61	0.25
31	*p*-cresol	3.00	0.47	0.82	1.61	0.23
32	benzyl alcohol (α-hydroxytoluene)	***2.98***	***0.25***	***0.80***	***1.60***	***0.38***
33	1-naphthol	4.10	0.13	1.52	1.72	0.27
34	2-naphthol	4.15	0.16	1.52	1.72	0.29
35	2-methoxyphenol (guaiacol)	3.10	0.41	0.84	0.62	0.38
36	3-methoxyphenol	3.28	0.84	0.88	1.67	0.28
37	4-methoxyphenol	3.23	0.77	0.90	1.61	0.35
		Aldehydes				
38	acetaldehyde	0.57	0.60	0.21	0.00	0.33
39	propionaldehyde	1.18	0.58	0.20	0.00	0.33
40	butyraldehyde	1.64	0.59	0.19	0.00	0.33
41	isobutyraldehyde	1.55	0.58	0.14	0.00	0.33
42	pentanal	***2.10***	***0.61***	***0.23***	***−0.06***	***0.30***
43	hexanal	***2.69***	***0.62***	***0.17***	***−0.03***	***0.30***
44	heptanal	3.28	0.63	0.14	0.00	0.33
45	octanal	3.76	0.61	0.16	0.00	0.33
46	nonanal	4.26	0.62	0.15	0.00	0.33
47	2-propenal	0.87	0.58	0.32	0.00	0.33
48	*trans*-but-2-ene-1-al	1.68	0.63	0.39	0.00	0.36
49	2-methylpropenal	1.34	0.46	0.40	0.00	0.36
50	benzaldehyde	2.63	0.63	0.82	0.00	0.28
51	paraldehyde	2.57	0.56	0.14	0.00	0.50
		Ketones				
52	propanone	1.05	0.65	0.18	0.11	0.36
53	2-butanone	***1.43***	***0.70***	***0.29***	***−0.14***	***0.33***
54	2-pentanone	***2.19***	***0.69***	***0.08***	***−0.06***	***0.34***
55	2-hexanone	***2.62***	***0.69***	***0.13***	***−0.08***	***0.35***
56	2-heptanone	***3.20***	***0.70***	***0.06***	***−0.06***	***0.36***
57	2-octanone	3.68	0.67	0.11	0.00	0.37
58	2-nonanone	4.15	0.66	0.12	0.00	0.37
59	2-decanone	4.67	0.67	0.11	0.00	0.37
60	2-undecanone	5.16	0.68	0.10	0.00	0.37
61	2-dodecanone	5.60	0.68	0.10	0.00	0.37
62	cyclopentanone	***2.25***	***0.66***	***0.42***	***−0.08***	***0.46***
63	cyclohexanone	***2.73***	***0.66***	***0.47***	***−0.13***	***0.48***
64	cycloheptanone	3.41	0.64	0.44	0.00	0.41
65	acetophenone	3.12	0.60	0.82	0.00	0.35
66	ethylphenyl ketone	3.63	0.50	0.80	0.00	0.37

Table A1 (Continued)

number	compounds	δ_4	ω_4	ϵ_4	α_4	β_4
		Ethers				
67	diethyl ether	1.73	0.10	0.04	0.00	0.33
68	dipropyl ether	***2.67***	***0.11***	***0.00***	***0.05***	***0.13***
69	dibutyl ether	***3.73***	***0.11***	***−0.06***	***0.06***	***0.13***
70	furan	1.10	0.43	0.37	0.00	0.09
71	benzofuran	3.00	0.44	0.89	0.00	0.11
72	tetrahydrofuran	***1.76***	***0.22***	***0.43***	***−0.09***	***0.33***
73	2-methyltetrahydrofuran	2.23	0.24	0.24	0.00	0.39
74	1,4-dioxane	***2.00***	***0.34***	***0.43***	***0.03***	***0.37***
75	1,4-benzodioxan	3.51	0.71	0.87	0.00	0.25
76	anisole (methoxybenzene)	***2.80***	***0.50***	***0.72***	***−0.03***	***0.11***
77	phenetole (ethylphenyl ether)	***3.32***	***0.46***	***0.60***	***−0.01***	***0.14***
		Nitrogen Compounds				
78	nitromethane	0.99	1.02	0.31	0.17	0.23
79	nitroethane	***1.40***	***1.17***	***0.34***	***0.17***	***0.17***
80	1-nitropropane	***2.05***	***1.13***	***0.22***	***0.15***	***0.17***
81	1-nitrobutane	***2.58***	***1.11***	***0.21***	***0.14***	***0.18***
82	1-nitropentane	***3.06***	***1.11***	***0.23***	***0.13***	***0.18***
83	1-nitrohexane	3.61	1.12	0.20	0.00	0.21
84	nitrocyclohexane	3.80	0.94	0.44	0.00	0.23
85	nitrobenzene	***3.11***	***1.04***	***0.86***	***0.00***	***0.18***
86	2-nitrotoluene	3.40	0.82	0.87	0.00	0.20
87	3-nitrotoluene	3.62	0.81	0.87	0.00	0.18
88	4-nitrotoluene	3.67	0.81	0.87	0.00	0.20
89	2-nitrophenol	3.18	0.55	1.02	0.14	0.27
90	3-nitrophenol	3.80	1.39	1.05	2.23	0.17
91	4-nitrophenol	3.89	1.58	1.07	2.32	0.19
92	cyanomethane (acetonitrile)	0.93	1.00	0.24	0.20	0.23
93	cyanoethane (proprionitrile)	1.34	1.04	0.16	0.06	0.26
94	1-cyanopropane (butyronitrile)	***1.64***	***1.10***	***0.23***	***0.00***	***0.32***
95	1-cyanobutane (valeronitrile)	***2.26***	***1.09***	***0.16***	***0.02***	***0.34***
96	1-cyanopentane (hexanenitrile)	***2.78***	***1.09***	***0.15***	***0.03***	***0.34***
97	phenyl cyanide (benzonitrile)	2.67	0.89	0.74	0.00	0.24
98	1-cyanohexane (heptanenitrile)	3.35	1.04	0.16	0.00	0.26
99	1-cyanoheptane (caprylonitrile)	3.85	1.04	0.16	0.00	0.26
100	1-cyanooctane (nonanenitrile)	4.23	1.04	0.16	0.00	0.26
101	1-cyanononane (decanenitrile)	4.73	1.05	0.16	0.00	0.26
102	1-cyanodecane (hendecanenitrile)	5.21	1.05	0.15	0.00	0.26
103	2-cyanophenol	2.89	1.08	0.92	2.09	0.24
104	3-cyanophenol	3.41	1.43	0.93	2.18	0.20
105	4-cyanophenol	3.60	1.54	0.94	2.23	0.21
106	ammonia	0.25	0.08	0.14	0.40	0.45
107	methylamine	0.78	0.01	0.25	0.45	0.42
108	ethylamine (aminoethane)	1.16	0.01	0.24	0.45	0.44
109	*n*-propylamine	1.64	0.02	0.23	0.45	0.44
110	*n*-butylamine	2.12	0.02	0.22	0.45	0.44
111	*n*-pentylamine	2.65	0.03	0.21	0.45	0.44
112	*n*-hexylamine	3.18	0.04	0.20	0.45	0.44
113	*n*-octylamine	4.05	0.05	0.19	0.45	0.44
114	dimethylamine	1.16	−0.06	0.19	0.23	0.48
115	diethylamine	1.98	−0.04	0.15	0.23	0.50
116	di-*n*-propylamine	2.97	−0.02	0.12	0.23	0.50
117	di-*n*-butylamine	3.98	0.00	0.11	0.23	0.50
118	di-*n*-pentylamine	4.21	0.00	0.10	0.23	0.50
119	trimethylamine	1.27	−0.17	0.14	0.00	0.49
120	triethylamine	2.75	−0.28	0.10	0.00	0.58
121	benzylamine (α-amino-toluene)	2.99	0.26	0.83	0.28	0.52
122	piperidine	2.57	−0.02	0.42	0.28	0.50
123	*N*-methylpiperidine	2.72	−0.05	0.32	0.00	0.51
124	*N*-ethylpiperidine	3.18	−0.10	0.30	0.00	0.46
125	pyrrole	1.81	0.44	0.61	1.16	0.21
126	*N*-methylpyrrole	1.89	0.57	0.56	0.00	0.23
127	pyridine	***1.92***	***0.31***	***0.66***	***0.00***	***0.53***
128	2-methylpyridine (2-picoline)	***2.42***	***0.18***	***0.60***	***−0.03***	***0.58***
129	3-methylpyridine (3-picoline)	***2.48***	***0.31***	***0.67***	***−0.08***	***0.64***
130	4-methylpyridine (4-picoline)	***2.37***	***0.31***	***0.74***	***−0.11***	***0.66***
131	2,3-dimethylpyridine (2,3-lutidine)	***2.89***	***0.21***	***0.71***	***−0.13***	***0.64***
132	2,4-dimethylpyridine (2,4-lutidine)	***2.89***	***0.22***	***0.65***	***−0.15***	***0.68***
133	2,5-dimethylpyridine (2,5-lutidine)	***2.87***	***0.19***	***0.68***	***−0.15***	***0.66***

Table A1 (Continued)

number	compounds	δ_4	ω_4	ϵ_4	α_4	β_4
134	2,6-dimethylpyridine (2,6-lutidine)	***2.86***	***0.11***	***0.55***	***−0.04***	***0.59***
135	3,4-dimethylpyridine (3,4-lutidine)	***2.96***	***0.34***	***0.79***	***−0.19***	***0.79***
136	3,5-dimethylpyridine (3,5-lutidine)	***2.96***	***0.32***	***0.73***	***−0.17***	***0.74***
137	pyrazine	1.74	0.59	0.63	0.00	0.45
138	2-methylpyrazine	2.10	0.50	0.63	0.00	0.47
139	pyrimidine	1.65	0.67	0.61	0.00	0.47
140	3-chloropyridine	***2.63***	***0.30***	***0.76***	***0.14***	***0.32***
141	*N*,*N*-dimethylformamide	2.03	1.29	0.37	0.00	0.54
142	*N*,*N*-dimethylacetamide	2.57	1.30	0.36	0.00	0.57
143	benzamide	3.97	1.10	0.99	1.38	0.49
	Carboxylic Acids					
144	acetic acid	1.05	0.53	0.27	1.72	0.32
145	propanoic acid	1.62	0.55	0.23	1.70	0.33
146	butanoic acid	2.20	0.53	0.21	1.70	0.33
147	pentanoic acid	2.76	0.50	0.21	1.70	0.33
148	hexanoic acid	3.33	0.53	0.17	1.70	0.33
149	heptanoic acid	3.89	0.55	0.15	1.70	0.33
150	octanoic acid	4.43	0.55	0.15	1.70	0.33
151	nonanoic acid	5.00	0.56	0.13	1.70	0.33
152	decanoic acid	5.54	0.57	0.12	1.70	0.33
153	undecanoic acid	6.12	0.59	0.10	1.70	0.33
154	dodecanoic acid	6.67	0.60	0.08	1.70	0.33
	Esters					
155	methyl formate	0.64	0.67	0.19	0.00	0.28
156	ethyl formate	1.25	0.68	0.15	0.00	0.28
157	*n*-propyl formate	1.86	0.64	0.13	0.00	0.28
158	*n*-butyl formate	2.40	0.65	0.12	0.00	0.28
159	*n*-pentyl formate	2.95	0.67	0.10	0.00	0.28
160	methyl acetate	1.33	0.61	0.14	0.00	0.33
161	ethyl acetate	1.77	0.61	0.11	0.00	0.33
162	*n*-propyl acetate	***2.40***	***0.55***	***0.00***	***0.02***	***0.31***
163	*n*-butyl acetate	***2.93***	***0.54***	***−0.03***	***0.02***	***0.32***
164	*n*-pentyl acetate	***3.53***	***0.54***	***−0.10***	***0.06***	***0.32***
165	*n*-hexyl acetate	3.87	0.62	0.06	0.00	0.33
166	*n*-heptyl acetate	4.39	0.63	0.05	0.00	0.33
167	*n*-octyl acetate	4.90	0.65	0.03	0.00	0.33
168	vinyl acetate	1.50	0.56	0.22	0.00	0.31
169	ailyi acetate	2.05	0.67	0.20	0.00	0.36
170	phenyl acetate	3.11	0.88	0.66	0.00	0.39
171	methyl benzoate	3.48	0.43	0.73	0.00	0.33
172	ethyl benzoate	3.89	0.47	0.69	0.00	0.33
173	*n*-propyl benzoate	4.57	0.41	0.68	0.00	0.33
174	*n*-butyl benzoate	5.07	0.41	0.67	0.00	0.33
175	dimethyl phthalate	4.49	1.02	0.78	0.00	0.64
	Halogen Compounds					
176	fluoroethane	0.21	0.44	0.05	0.00	0.07
177	1-fluoropropane	0.74	0.45	0.03	0.00	0.07
178	1-fluorohexane	2.65	0.48	0.00	0.00	0.07
179	1-fluorooctane	3.57	0.50	−0.02	0.00	0.07
180	1,1,1-trifluorooctane	***3.61***	***0.77***	***−0.56***	***−0.02***	***0.00***
181	fluorobenzene	***2.08***	***0.43***	***0.40***	***0.07***	***0.04***
182	hexafluorobenzene	***2.15***	***0.87***	***−0.14***	***−0.10***	***−0.06***
183	trifluoromethylbenzene	***2.44***	***0.75***	***0.05***	***0.04***	***0.00***
184	2.2,2-trifluoroethanol	0.78	0.77	0.02	1.61	0.18
185	hexafluoropropan-1-oi	1.20	0.98	−0.24	2.18	0.07
186	dodecafluoroheptan-1-ol	3.28	1.18	−0.64	1.84	0.16
187	2-fluorophenol	2.38	0.36	0.66	1.72	0.19
188	3-fluorophenol	2.61	0.84	0.67	1.92	0.12
189	4-fluorophenol	2.61	0.79	0.67	1.78	0.17
190	methoxyflurane [ethane,2,2-dichloro-1,1-difluoro-1-methoxy-]	2.30	0.85	0.11	0.20	0.10
191	isoflurane [ethane,2-chloro-2-(difluoromethoxy)-1,1,1-trifluoro-]	1.41	0.88	−0.24	0.28	0.12
192	enflurane [ethane,2-chloro-1-(difluoromethoxy)-1,1,2-trifluoro-]	1.63	0.73	−0.23	0.34	0.09
193	halothane (2-bromo-2-chloro-1,1,1-trifluoroethane)	1.77	0.47	0.10	0.42	0.04
194	dichloromethane	1.25	0.52	0.39	0.28	0.04
195	trichloromethane	***1.50***	***0.15***	***0.66***	***0.63***	***0.03***
196	tetrachloromethane	***2.05***	***0.01***	***0.57***	***0.10***	***−0.01***
197	chloroethane	1.15	0.37	0.23	0.00	0.07
198	1,2-dichloroethane	1.75	0.57	0.42	0.28	0.08

Table A1 (Continued)

number	compounds	δ_4	ω_4	ϵ_4	α_4	β_4
199	1,1,1-trichloroethane	2.07	0.27	0.37	0.00	0.07
200	1,1,2-trichloroethane	2.37	0.57	0.50	0.37	0.06
201	1-chloropropane	1.68	0.37	0.22	0.00	0.07
202	2-chloropropane	1.51	0.32	0.18	0.00	0.09
203	1-chlorobutane	***2.11***	***0.32***	***0.31***	***0.04***	***0.05***
204	1-chloropentane	***2.60***	***0.32***	***0.33***	***0.02***	***0.06***
205	1-chlorohexane	***3.14***	***0.32***	***0.30***	***0.03***	***0.07***
206	chlorocyclohexane	3.25	0.30	0.45	0.00	0.07
207	1-chloroheptane	3.78	0.39	0.19	0.00	0.07
208	1-chlorooctane	4.27	0.40	0.19	0.00	0.07
209	1,1-dichloroctane	1.65	0.42	0.32	0.28	0.07
210	1,1-dichloroethene	1.49	0.19	0.36	0.00	0.04
211	*cis*-1,2-dichloroethene	1.61	0.54	0.44	0.31	0.04
212	*trans*-1,2-dichloroethene	1.56	0.24	0.43	0.25	0.04
213	chlorobenzene	***2.62***	***0.32***	***0.75***	***0.04***	***0.06***
214	1,2-dichlorobenzene	3.21	0.44	0.87	0.00	0.03
215	1,3-dichlorobenzene	3.15	0.39	0.85	0.00	0.01
216	1,4-dichlorobenzene	3.18	0.44	0.83	0.00	0.01
217	benzyl chloride (2-chlorotoluene)	3.10	0.38	0.82	0.00	0.24
218	2-chlorophenol	2.83	0.46	0.85	0.90	0.23
219	3-chlorophenol	3.28	0.77	0.91	1.95	0.11
220	4-chlorophenol	3.27	0.77	0.92	1.89	0.15
221	bromomethane	0.93	0.27	0.40	0.00	0.07
222	dibromomethane	1.78	0.37	0.71	0.28	0.07
223	tribromomethane	2.44	0.19	0.97	0.42	0.04
224	bromoethane (ethylbromide)	1.46	0.24	0.37	0.00	0.09
225	1-bromopropane	***1.95***	***0.25***	***0.40***	***0.08***	***0.07***
226	1-bromobutane	***2.40***	***0.25***	***0.44***	***0.04***	***0.08***
227	1-bromopentane	***2.99***	***0.26***	***0.38***	***0.06***	***0.08***
228	bromocyclohexane	3.44	0.22	0.62	0.00	0.12
229	1-bromohexane	3.49	0.25	0.35	0.00	0.09
230	1-bromoheptane	4.03	0.26	0.34	0.00	0.09
231	1-bromooctane	4.46	0.26	0.34	0.00	0.09
232	1-bromononane	4.93	0.26	0.34	0.00	0.09
233	bromobenzene	***2.92***	***0.26***	***0.88***	***0.07***	***0.08***
234	benzyl bromide	3.13	0.54	1.01	0.00	0.15
235	1,2-dibromobenzene	3.77	0.44	1.19	0.00	0.03
236	1,3-dibromobenzene	3.70	0.34	1.17	0.00	0.03
237	1,4-dibromobenzene	3.72	0.33	1.15	0.00	0.03
238	2-bromophenol	3.01	0.34	1.04	0.99	0.23
239	3-bromophenol	3.47	0.78	1.06	1.98	0.12
240	4-bromophenol	3.43	0.77	1.08	1.89	0.15
241	iodomethane	1.16	0.02	0.68	0.00	0.09
242	diiodomethane	2.08	−0.29	1.45	0.14	0.17
243	iodoethane	1.67	−0.01	0.64	0.00	0.11
244	1-iodopropane	2.24	0.00	0.63	0.00	0.11
245	1-iodobutane	2.74	0.00	0.63	0.00	0.11
246	1-iodopentane	3.25	0.01	0.62	0.00	0.11
247	1-iodohexane	3.74	0.01	0.62	0.00	0.11
248	iodobenzene	2.89	0.19	1.19	0.00	0.09
249	2-iodophenol	3.10	0.20	1.36	1.13	0.25
250	3-iodophenol	3.55	0.58	1.37	1.98	0.13
251	4-iodophenol	3.49	0.60	1.38	1.92	0.15
		Sulfur Compounds				
252	ethylthiol	1.52	0.08	0.39	0.00	0.17
253	ailyi thiol (2-propene-1-thiol)	1.84	0.06	0.54	0.00	0.15
254	1-propanethiol	2.04	0.08	0.39	0.00	0.17
255	1-butanethiol	***2.45***	***0.09***	***0.39***	***0.20***	***0.07***
256	1-pentanethiol	***2.97***	***0.09***	***0.38***	***0.20***	***0.08***
257	*n*-hexanethiol	***3.48***	***0.10***	***0.36***	***0.19***	***0.09***
258	1-heptanethiol	4.01	0.11	0.36	0.00	0.17
259	1-octanethiol	4.65	0.11	0.35	0.00	0.17
260	1-nonanethiol	5.18	0.11	0.35	0.00	0.17
261	1-decanethiol	5.71	0.12	0.34	0.00	0.17
262	thiophene	***1.77***	***0.23***	***0.73***	***0.09***	***0.07***
263	thiophenol (benzenethiol)	2.68	0.30	1.00	0.25	0.12
264	benzo(*b*)thiophene	3.41	0.13	1.32	0.00	0.15
265	dibenzothiophen	5.01	0.26	1.96	0.00	0.13
266	dimethyl sulfide (methylthiomethane)	1.55	0.08	0.40	0.00	0.21

Table A1 (Continued)

number	compounds	δ_4	ω_4	ϵ_4	α_4	β_4
267	diethyl sulfide	2.45	0.09	0.37	0.00	0.23
268	di-*n*-propyl sulfide	3.48	0.11	0.36	0.00	0.23
269	di-*n*-butyl sulfide	4.32	0.12	0.35	0.00	0.23
270	sulfur hexafluoride	0.41	0.20	−0.60	0.00	0.00
271	carbon disulfide	1.34	−0.45	0.88	0.00	0.05
		Hydrocarbons				
272	ethene	0.03	0.03	0.11	0.00	0.05
273	propene	0.70	0.00	0.10	0.00	0.05
274	but-1-ene	1.24	0.00	0.10	0.00	0.05
275	buta-1,3-diene	1.02	0.03	0.32	0.00	0.07
276	cyclopentene	1.88	−0.03	0.34	0.00	0.07
277	pent-1-ene	1.81	0.01	0.09	0.00	0.05
278	2-methylbuta-1,3-diene	1.58	0.03	0.31	0.00	0.07
279	1-hexene	***2.24***	***0.06***	***0.10***	***0.03***	***0.00***
280	2,3-dimethylbuta-1,3-diene	2.14	−0.02	0.35	0.00	0.10
281	1-methylcyclopentene	2.35	−0.03	0.33	0.00	0.07
282	cyclohexene	***2.62***	***−0.07***	***0.27***	***0.08***	***0.04***
283	1,4 cyclohexadiene	***2.25***	***0.02***	***0.58***	***0.03***	***0.06***
284	1,3 cyclohexadiene	***2.32***	***0.05***	***0.40***	***0.06***	***0.06***
285	1-methylcyclohexene	2.91	−0.08	0.39	0.00	0.07
286	cycloheptene	3.02	−0.07	0.41	0.00	0.07
287	1-heptene	***2.86***	***0.07***	***0.01***	***0.06***	***0.01***
288	1-methylcycloheptene	3.34	−0.08	0.43	0.00	0.07
289	1-octene	***3.25***	***0.06***	***0.08***	***0.03***	***0.01***
290	1-nonene	***3.79***	***0.07***	***0.04***	***0.04***	***0.01***
291	1-decene	***4.27***	***0.07***	***0.06***	***0.04***	***0.01***
292	undec-1-ene	4.78	0.01	0.09	0.00	0.05
293	dodec-1-ene	5.28	0.01	0.09	0.00	0.05
294	ethyne	−0.27	0.14	0.19	0.59	0.11
295	propyne	0.61	0.15	0.18	0.34	0.11
296	but-1-yne	1.12	0.12	0.18	0.34	0.11
297	1-pentyne	1.62	0.14	0.17	0.34	0.09
298	1-hexyne	***2.10***	***0.21***	***0.16***	***0.22***	***0.03***
299	2-hexyne	***2.75***	***0.11***	***−0.09***	***0.24***	***0.07***
300	3-hexyne	***2.57***	***0.04***	***−0.02***	***0.27***	***0.06***
301	1-heptyne	***2.74***	***0.21***	***0.06***	***0.26***	***0.04***
302	1-octyne	***3.22***	***0.22***	***0.07***	***0.25***	***0.04***
303	oct-2-yne	3.37	0.21	0.23	0.00	0.07
304	4-octyne	***3.55***	***0.07***	***−0.06***	***0.23***	***0.05***
305	1-nonyne	***3.71***	***0.22***	***0.07***	***0.25***	***0.04***
306	1-decyne	***4.18***	***0.22***	***0.09***	***0.24***	***0.04***
307	dodec-1-yne	5.30	0.19	0.13	0.34	0.07
308	benzene	***1.95***	***0.26***	***0.57***	***−0.04***	***0.08***
309	toluene	***2.41***	***0.29***	***0.62***	***−0.10***	***0.08***
310	ethylbenzene	***2.92***	***0.26***	***0.58***	***−0.07***	***0.08***
311	styrene (vinylbenzene)	2.64	0.19	0.85	0.00	0.12
312	phenylethyne (phenylacetylene)	2.77	−0.12	0.68	0.34	0.17
313	*o*-xylene (1,2-dimethylbenzene)	2.93	0.21	0.66	0.00	0.12
314	*m*-xylene	2.89	0.18	0.62	0.00	0.12
315	*p*-xylene	2.90	0.19	0.61	0.00	0.12
316	*n*-propylbenzene	3.31	0.18	0.60	0.00	0.11
317	allyl benzene	3.06	0.20	0.72	0.00	0.16
318	1,3,5-trimethylbenzene	3.37	0.15	0.65	0.00	0.14
319	*n*-butylbenzene	3.81	0.19	0.60	0.00	0.11
320	pentamethylbenzene	4.57	0.19	0.85	0.00	0.15
321	hexamethylbenzene	5.21	0.19	0.95	0.00	0.15
322	diphenylmethane	4.55	0.41	1.22	0.00	0.20
323	biphenyl	4.16	0.25	1.36	0.00	0.16
324	naphthalene	***3.54***	***0.40***	***1.28***	***−0.20***	***0.21***
325	azulene	***3.84***	***0.35***	***1.47***	***−0.08***	***0.19***
326	anthracene	4.69	−0.02	2.29	0.00	0.19
327	phenanthrene	4.99	0.10	2.06	0.00	0.19
328	cyclopentane	***2.34***	***−0.11***	***0.07***	***0.11***	***0.01***
329	pentane	***1.98***	***0.00***	***0.00***	***0.00***	***0.00***
330	methylcyclopentane	2.45	−0.04	0.23	0.00	0.00
331	cyclohexane	***2.73***	***−0.11***	***0.16***	***0.06***	***0.02***
332	hexane	***2.47***	***0.00***	***0.00***	***0.00***	***0.00***
333	2,2-dimethylbutane	***2.22***	***0.05***	***−0.04***	***−0.02***	***0.00***
334	2,3-dimethylbutane	***2.28***	***0.04***	***0.03***	***−0.03***	***0.00***

Table A1 (Continued)

number	compounds	δ_4	ω_4	ϵ_4	α_4	β_4
335	methylcyclohexane	***3.00***	***−0.07***	***0.19***	***0.01***	***0.03***
336	cycloheptane	***3.15***	***−0.14***	***0.38***	***0.01***	***0.04***
337	2,4-dimethylpentane	***2.53***	***0.06***	***0.05***	***−0.06***	***−0.01***
338	heptane	***2.96***	***0.00***	***0.00***	***0.00***	***0.00***
339	2,2-dimethylpentane	***2.71***	***0.07***	***−0.09***	***−0.01***	***0.00***
340	2,3-dimethylpentane	***2.86***	***0.03***	***0.00***	***−0.01***	***0.00***
341	2,2,3-trimethylbutane	***2.44***	***0.08***	***0.18***	***−0.13***	***0.02***
342	methylcycloheptane	3.60	−0.10	0.30	0.00	0.00
343	cyclooctane	***3.52***	***−0.15***	***0.56***	***−0.04***	***0.05***
344	octane	***3.46***	***0.00***	***0.00***	***0.00***	***0.00***
345	2,2,4-trimethylpentane	***3.26***	***0.10***	***−0.25***	***0.00***	***0.00***
346	2,3-dimethylhexane	***3.33***	***0.04***	***−0.04***	***0.00***	***0.00***
347	2,4-dimethylhexane	***3.15***	***0.06***	***−0.03***	***−0.04***	***0.00***
348	3,4-dimethylhexane	***3.10***	***0.03***	***0.18***	***−0.07***	***0.00***
349	2,3,4-trimethylpentane	***3.33***	***0.04***	***−0.04***	***−0.01***	***0.01***
350	*cis*-1,2-dimethylcyclohexane	***3.32***	***−0.07***	***0.35***	***−0.05***	***0.04***
351	*trans*-1,2 dimethylcyclohexane	***3.30***	***−0.04***	***0.25***	***−0.04***	***0.03***
352	*cis*-1,4-dimethylcyclohexane	***3.22***	***−0.04***	***0.31***	***−0.05***	***0.02***
353	*trans*-1,4 dimethylcyclohexane	***3.10***	***−0.02***	***0.32***	***−0.07***	***0.02***
354	*trans*-hydrindane	***3.67***	***−0.12***	***0.58***	***−0.06***	***0.03***
355	*cis*-hydrindane	***3.70***	***−0.15***	***0.68***	***−0.06***	***0.05***
356	nonane	***3.95***	***0.00***	***0.00***	***0.00***	***0.00***
357	adamantane	***4.10***	***−0.24***	***0.77***	***−0.04***	***0.11***
358	*trans*-decalin	***4.05***	***−0.13***	***0.70***	***−0.10***	***0.06***
359	*cis*-decalin	***4.16***	***−0.15***	***0.78***	***−0.10***	***0.05***
360	cyclodecane	***4.35***	***−0.15***	***0.73***	***−0.07***	***0.04***
361	decane	***4.45***	***0.00***	***0.00***	***0.00***	***0.00***
362	undecane	***4.94***	***0.00***	***0.00***	***0.00***	***0.00***
363	dodecane	***5.43***	***0.00***	***0.00***	***0.00***	***0.00***
364	tridecane	***5.93***	***0.00***	***0.00***	***0.00***	***0.00***
365	tetradecane	***6.42***	***0.00***	***0.00***	***0.00***	***0.00***
		Miscellaneous				
366	hexamethyldisilane	***3.80***	***0.17***	***−0.69***	***0.14***	***−0.02***
367	hexamethyldisiloxane	***3.75***	***0.24***	***−0.99***	***0.24***	***−0.03***
368	tetramethyltin	***2.78***	***0.07***	***−0.23***	***0.12***	***0.01***
369	triethyl phosphate	4.10	0.95	0.00	0.00	0.77

[a] The parameters are designated as dispersion (δ_4), orientation (ω_4), polarizability-induction (ϵ_4), acidity (α_4), and basicity (β_4). Data italicized and bold-faced are from Laffort et al.[1], and the other data are from Abraham,[6] adapted according to eq 1.

4. Other Predictive Studies. In recent publications devoted to evaluating the solvation parameters from the molecular structure, a first group of approaches is based on a rather great number of molecular fragments and physicochemical properties characterizing each compound, to which are applied LMRA or NN (Platts et al.,[19,21] Havelec and Sevcik,[20] and Jover et al.[22]). Other approaches, perhaps more promising, are related to theoretical chemistry (Svozil and Sevcik,[23] Daerden and Ghafourian,[24,25] Platts,[26,27] Lamarche et al.,[28] and Oliferenko et al.[29]). A comparative study limited to the parameters of acidity and basicity in the present work and in the most recent publications is reported in Table 3. To have a homogeneity of presentation, the experimental solvation parameters of acidity and basicity have to all be respectively expressed as α_2 and β_2, according to eq 1.

In the limited comparisons reported in Table 3, the performances of the three approaches are very similar, with, however, a trend for ours to be slightly better for the acidity parameter and slightly worse for the basicity parameter.

GENERAL CONCLUSION

The compared performances of various theoretical and empirical approaches proposed until now, including the present work, to establish the solvation parameters from the molecular structure, deserve deeper, further study in order to optimize, possibly via combined procedures, both the reliability of the obtained values and, if possible, the ease to do that. (The SMT program will be soon available on the Web site http://paul.laffort.free.fr.)

ACKNOWLEDGMENT

The authors sincerely thank Dominique Valentin (Centre des Sciences du Goût of Dijon) and an anonymous reviewer for their help in the final version of this paper. They are also grateful to Anna Montmayeur for her editing assistance.

Note Added after ASAP Publication. This article was released ASAP on June 8, 2006, with an erroneous reference in the second paragraph of the Discussion. The correct version was posted on June 15, 2006.

APPENDIX A

Table A1 consists of the learning experimental data set of solvation parameters for the 369 solutes used in this work.

REFERENCES AND NOTES

(1) Laffort, P.; Chauvin, F.; Dallos, A.; Callegari, P.; Valentin, D. Solvation parameters. Part 1: Mutual improvements of several

approaches and determination of two first sets of optimized values. *J. Chromatogr., A* **2005**, *1100*, 90−107.
(2) Abraham, M. H.; Whiting, G.; Doherty, R. M.; Shuely, W. H. Hydrogen bonding. Part 13. A new method for the characterisation of GLC stationary phases-The Laffort data set. *J. Chem. Soc., Perkin Trans.* **1990**, *2*, 1451−1460.
(3) Laffort, P.; Patte F. The solubility factors in gas−liquid chromatography. Comparison between two approaches and application to some biological studies. *J. Chromatogr*. **1976**, *126*, 625−639.
(4) Patte, F.; Etcheto, M.; Laffort, P. Solubility factors for 240 solutes and 207 stationary phases in Gas−Liquid Chromatography. *Anal. Chem*. **1982**, *54*, 2239−2247.
(5) Laffort, P.; Patte, F. Solubility factors established by gas−liquid chromatography: A balance-sheet. *J. Chromatogr*. **1987**, *406*, 51−74.
(6) Abraham, M. H. Scales of solute hydrogen-bonding: Their construction and application to physicochemical and biochemical processes. *Chem. Soc. Rev*. **1993**, *22*, 73−83.
(7) Dubois, J. E. French national policy for chemical information and the DARC system as a potential tool of this policy. *J. Chem. Doc*. **1973**, *13*, 8−13.
(8) Dubois, J. E.; Chretien, J. R.; Sojác, L.; Rijks, J. A. Topological analysis of the behaviour of linear alkenes up to tetradecenes in gas−liquid chromatography on squalane. *J. Chromatogr*. **1980**, *194*, 21−134.
(9) Dubois, J. E.; Loukianoff, M. DARC 'logic' method for the molar volume prediction. *SAR-QSAR Environ. Res*. **1993**, *1*, 63−75.
(10) Dubois, J. E.; Loukianoff, M.; Mercier, C. Topology and the quest for structural knowledge. *J. Chim. Phys*. **1992**, *89*, 1493−1506.
(11) Laffort, P. In *Odors and Deodorization in the Environment*; Martin, G., Laffort, P., Eds.; VCH Publications: New York, 1994; Chapter 6, pp 143−183.
(12) Chauvin, F.; Laffort, P. Determination of a polarizability index by means of molecular topology. *J. Chim. Phys*. **1997**, *94*, 1216−1233.
(13) Zamora, A. An algorithm for finding the smallest set of smallest rings. *J. Chem. Inf. Comput. Sci*. **1976**, *16*, 40−43.
(14) ChemAxon Ltd., 2005. http://www.chemaxon.com/marvin/ (accessed Dec 2005).
(15) Palm, K.; Stenberg, P.; Luthman, K.; Artursson, P. Polar molecular surface properties predict the intestinal absorption of drugs in humans. *Pharm. Res*. **1997**, *14*, 568−571.
(16) Ertl, P.; Rohde, B.; Selzer, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* **2000**, *43*, 3714−3717.
(17) Zissimos, A. M.; Abraham, M. H.; Klamt, A.; Eckert F.; Wood, J. A comparison between the two general sets of linear energy descriptors of Abraham and Klamt. *J. Chem. Inf. Comput. Sci*. **2002**, *42*, 1320−1331.
(18) Abraham, M. H.; McGowan, J. C. The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography. *Chromatographia* **1987**, *23*, 243−246.
(19) Platts, J. A.; Butina, D.; Abraham, M. H.; Hersey, A. Estimation of molecular linear free energy relation descriptors using a group contribution approach. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 835−845.
(20) Havelec, P.; Sevcik, J. G. K. Extended additivity model of parameter log (L16). *J. Phys. Chem. Ref. Data* **1996**, *25*, 1483−1439.
(21) Platts, J. A.; Butina, D.; Abraham, M. H.; Hersey, A. Estimation of molecular linear free energy relation descriptors using a group contribution approach. 2. Prediction of partition coefficients. *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 71−80.
(22) Jover, J.; Bosque, R.; Sales, J. Determination of Abraham solute parameters from molecular structure. *J. Chem. Inf. Comput. Sci*. **2004**, *44*, 1098−1106.
(23) Svozil, D.; Sevcik, J. G. K. Neural network prediction of the solvatochromic polarity/polarisability parameter π_2^H. *J. Chem. Inf. Comput. Sci*. **1997**, *37*, 338−342.
(24) Dearden, J. C.; Ghafourian, T. Hydrogen bonding parameters for QSAR: Comparison of indicator variables, hydrogen bond counts, molecular orbital and other parameters. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 231−235.
(25) Dearden, J. C.; Ghafourian, T. The use of atomic charges and orbital energies as hydrogen-bonding -donor parameters for QSAR studies: Comparison of MNDO, AM1 and PM3 Methods. *J. Pharm. Pharmacol*. **2000**, *52*, 603−610.
(26) Platts, A. J. Theoretical prediction of hydrogen bond donor capacity. *Phys. Chem. Chem. Phys*. **2000**, *2* (5), 973−980.
(27) Platts, A. J. Theoretical prediction of hydrogen bond basicity. *Phys. Chem. Chem. Phys*. **2000**, *2* (14), 3115−3120.
(28) Lamarche, O.; Platts, A. J.; Hersey, A. Theoretical Prediction of the polarity/ polarizability parameter π_2^H. *Phys. Chem. Chem. Phys.* **2001**, *3* (14), 2747−2753.
(29) Oliferenko, A. A.; Oliferenko, P. V.; Hudleston, J. G.; Rogers, R. D.; Palyulin, V. A.; Zefirov, N. S.; Katrizky, A. R. Theoretical scales of hydrogen bond acidity and basicity for application in QSAR/QSPPR studies and drug design. Partitioning of aliphatic compounds. *J. Chem. Inf. Comput. Sci*. **2002**, *42*, 1320−1331.

CI0600152