Solvation Parameters. Part 4: What is the Impact of Recent Chromatographic Data Update?

P. Laffort^{*} and P. Héricourt

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

Abstract: Accurate solvation parameters have been established using solely gas/liquid partition data by gas/chromatography from the Kováts group (Laffort *et al.* J. Chromatog. A, 2005, 1100, 90-107). Because these chromatographic data have been recently updated (Kováts and Kresz, J. Chromatog. A, 2006, 1113, 206-219), the solvation parameters of the involved solutes and stationary phases have been entirely computed on again. It appears from this enquiry that the solute parameters values (127 compounds) are almost identical, whereas solvent parameters (11 stationary phases) have been slightly improved. A simple method of stationary phases classification is derived from these results, as well as a method for further determination of solute parameters.

INTRODUCTION

The conceptual definition of solvation parameters or descriptors can be expressed as follows: if a matrix of a given solubility property can be expressed as a product of matrices A*B, then A and B are respectively matrices of solute and solvent solvation parameters. Because infinite number of matrices A and B can provide such product, some rules and procedures have been developed in order, among others, to get parameters mutually as independent as possible (i.e. faintly correlated), and clearly reflecting understandable physicochemical properties [1-3].

In addition to these rules and procedures, the second tool needed to rightly characterize the solvation parameters of solutes is a reliable experimental database of solubility properties. Recently, Laffort *et al.* [1] used an accurate matrix of retention indices in gas-liquid chromatography (GLC) for 133 solutes and 10 stationary phases. Eight of these phases, of polar nature, were established by Kováts and co-authors [4-7], and the other two non polar phases were published by Laffort *et al.* [1].

However, one year later, Kováts and Kresz [8] published updated values of most of the GLC retention indices previously published in the references [4-7]. This update has raised two questions, the first one should be firstly solved before getting an answer to the second one:

- Have the remaining GLC values published by Laffort *et al.* [1], and principally those on apolar phases, to be similarly updated?
- Have these updated results on polar and apolar phases, an impact on the definitions of solvation parameters of solutes and solvents?

The aim of the present study is to answer these two points and, if needed, specify some consequences.

CHECKING THE REMAINING GLC VALUES

The Original Database According to [1]

This experimental data matrix of GLC retention indices RI (133 solutes x 10 stationary phases) is reported in Supplementary Information (Table SI-1), exactly as it was published in [1]. It includes RI values on two families of branched stationary phases, with respectively four and six branches, as reported in Fig. (1) (we call these two families of phases 4B and 6B).

The eight polar phases, synthesized by Cloux *et al.* [9], are all of 4B type. The RI values reported are exactly in agreement with the references [4-7]. The nature of these phases is summarized in Table 1.

The two apolar phases from which the experimental RI values are reported in Table **SI-1** are hydrocarbons of 6B type (C67 and C103). This hydrocarbon family was firstly explored by Riedo *et al.* [10] and later developed by Défayes *et al.* [11]. The advantage of the alkane phases of 6B type over the 4B type is, for a given molecular weight, a lower melting point together with a similar shape. Strictly speaking, the 6B alkane family only concerns compounds with odd number of carbon atoms. However, RI values on a hypothetical C78 hydrocarbon phase of the same series were interpolated, using the following equation:

$$RI_{C78-6B} = 0.4035 RI_{C103-6B} + 0.5965 RI_{C67-6B}$$
(1)

Similarly, RI values on a hypothetical alkane of infinite molecular weight were extrapolated using the following equation:

$$RI_{C_{\infty}-6B} = 2.8611 RI_{C103-6B} - 1.8611 RI_{C67-6B}$$
(2)

These two equations are based on the fact, every thing else being equal, that the GLC retention indices are linearly proportional to the inverse of the molecular weight of the stationary phases [11-13].

The Corrected Database According to [8]

As already mentioned, the overlapping of the two databases in [1] and [8] is not complete: the most important

^{*}Address correspondence to this author at the CNRS, Centre des Sciences du Goût UMR 5170, 15 rue Hugues Picardet, F.21000 Dijon, France; E-mail: laffort@cesg.cnrs.fr

Solvation Parameters. Part 4

difference concerns the absence of GLC retention indices for an alkane phase of infinite molecular weight, in the publication by Kováts and Kresz [8]. RI values for the primary bromo phase are also missing in this cited article. Otherwise, are included RI values for a secondary alcohol of the 4B family, established by Dallos *et al.* [14]. Finally, the RI values for C78, present in both databases, are from the 4B family in one case, and from the 6B family in the other case.

	Polar Interacting Group(s)	Х	Y
PCI	Primary chloro	CH ₂ CI	CH ₂ CH ₃
PBr	Primary bromo	CH_2Br	CH ₂ CH ₃
MTF	Monotrifluoromethyl	CH ₂ CF ₃	CH ₂ CH ₃
TTF	Tetrakistrifluoromethyl	CH ₂ CF ₃	CH ₂ CF ₃
ТМО	Tetramethoxy	OCH ₃	OCH ₃
PCN	Primary cyano	CH ₂ CN	CH ₂ CH ₃
PSH	Primary thiol	CH ₂ SH	CH ₂ CH ₃
РОН	Primary alcohol	CH ₂ 0H	CH ₂ CH ₃

Table 1.Structural Details of the Polar Phases Studied in [1].See X and Y Meaning in Fig. (1, Left Side)



Fig. (1). Schematic representation of the two types of stationary phases synthesized and used by the Kováts group, and applied by Laffort *et al.* [1]. Experimental retention indices of 133 solutes on 10 of these phases are reported in Table **SI-1**.

The corrected database as reported in Table **SI-2** (Supplementary Information) is exactly in agreement with Kováts and Kresz [8], with the unique exception of the 2-Methyl-2-hexanol on Tetrakistrifluoromethyl phase (TTF). Indeed, a minimal transcription error from the references [5] to [8] has been observed in this case.

Comparison of the Two Databases

The differences of RI values in Tables **SI-1** and **SI-2** are reported in Table **2**, for the eight phases present in both tables (including the C78 phase in its two types 4B and 6B).

It is observed in 99% of cases (i.e. 1054 out of 1064) that the corrective increments characteristic of the solutes are constant for the eight phases. The exceptions are highlighted in Table 2. Among the polar phases, only one minimal exception is observed. Therefore, these corrective increments must also be applied to the primary bromo phase.

In spite of the different molecular structure of the two C78 phases, their comparison only reveals nine non zero differences, from which eight are equal or less than 1 index unit, which can be considered as faint. The ninth value, equal to 4 index units, concerns the tetramethylsilane. This difference can not be considered as negligible, but it concerns one of the five solutes out of the 133 under study which presents RI values smaller than 500 (i.e. less than for pentane). It is well known that at this level, the experimental accuracy of chromatographic measurements is lower. In the same way as in our last publication on this topic [3], we have therefore decided to discard these five solutes from the final experimental data matrix.

The results about the two C78 phases reported in Table **2** are very interesting, in the sense that C78-6B being established from C67 and C103 (eq. 1), the same corrections must be applied to these latter phases, and therefore also to a hypothetical alkane of infinite molecular weight (eq. 2).

Taking into account these facts, Table **SI-3** (Supplementary Information) includes the updated matrix of experimental retention indices as follows:

- The data of Table **SI-1** have been corrected using the not highlighted corrective increments of Table **2**.
- The data concerning the secondary alcohol phase SOH have been added, as it appears in Table **SI-2**. Because data on SOH are missing for the 1,1,1trifluorooctane, this solute has been discarded.
- The data concerning the five solutes with RI values smaller than 500 have been also discarded, as we saw. The names of these five solutes have been highlighted in Table **SI-1**.

The C78-6B values have been preferred to the C78-4B ones, in order to maintain a closer consistency between the C78 and Cinf phases.

Finally, the experimental matrix reported in Table SI-3 concerns 127 solutes and 11 stationary phases.

IMPACT ON THE SOLVATION PARAMETERS

The tool applied for checking a possible impact of the updated GLC retention indices above reported, on the characterization of solvation parameters, is the algorithm named MMA (as Multiplicative Matrix Analysis), already applied in [1].

This algorithm was firstly presented in 1976 [15] and detailed in [1]). It is a tool to test theories on the basis of experimental data, whenever products of matrices are involved. The Fig. (2) summarizes how the program works in a particular application of gas-liquid-chromatography (GLC), where the experimental matrix R is a product of the matrix A of parameters of solutes and a matrix B of parameters of solvents.

For a given matrix of experimental data R, a theoretical matrix A is placed in INPUT. The program runs using successive iterations, until the reconstruction of the experimental matrix in OUTPUT is optimal. The quality of this reconstruction (R compared to A * B) only depends on the number of input parameters A, not on their nature. By contrast, the output parameters A resemble to the input parameters A only when the latter are involved in the phenomenon under study.

Table 2.Differences of GLC Retention Indices According to [1] and [8] (Tables SI-1 and SI-2 in the Present Study), for 133 Solutes
and the Eight Stationary Phases Present in Both Tables

ID [1]	ID [8]	Solutes	C78	РОН	TTF	MTF	PCL	ТМО	PSH	PCN
1	76	1-Butanol	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9
2	84	2-Methyl-2-propanol	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9
3	77	1-Pentanol	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
4	85	2-Methyl-2-butanol	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
5	78	1-Hexanol	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
6	131	Cyclohexanol	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
7	86	2-Methyl-2-pentanol	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
8	79	1-Heptanol	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	87	2-Methyl-2-hexanol	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1
10	80	2-Butanol	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
11	81	2-Pentanol	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
12	82	2-Hexanol	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2
13	83	2-Heptanol	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1
14	159	2-Phenylethanol	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
15	158	Benzyl alcool	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
16	96	Pentanal	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
17	97	Hexanal	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
18	91	2-Butanone	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
19	92	2-Pentanone	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
20	129	Cyclopentanone	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
21	93	2-Hexanone	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
22	130	Cyclohexanone	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
23	94	2-Heptanone	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1
24	99	Dipropylether	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
25	100	Dibutylether	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
26	126	Tetrahydrofuran	0.3	0.3	0.3	0.3	0.4	0.3	0.3	0.3
27	127	1,4-Dioxane	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
28	160	Methyl phenyl ether (Anisole)	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
29	161	Phenetole	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
30	69	Nitroethane	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
31	70	1-Nitropropane	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4
32	71	1-Nitrobutane	-0.6	-0.6	-0.6	-0.6	-0.6	-0.6	-0.6	-0.6
33	72	1-Nitropentane	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3
34	157	1-Nitrobenzene	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
35	65	1-Cyanoethane	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3
36	66	1-Cyanopropane (Butyronitrile)	1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.7
37	67	1-Cyanobutane (Valeronitrile)	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
38	68	1-Cyanopentane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
39	116	Pyridine	2.2	2.2	2.2	2.2	2.2	2.2	2.2	2.2
40	117	2-Picoline	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
41	118	3-Picoline	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6
42	119	4-Picoline	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6
43	143	2,3-Lutidine	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6

	T	1	1						(Table	2) contd
ID [1]	ID [8]	Solutes	C78	РОН	TTF	MTF	PCL	ТМО	PSH	PCN
44	144	2,4-Lutidine	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
45	145	2,5-Lutidine	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
46	146	2,6-Lutidine	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
47	147	3,4-Lutidine	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8
48	148	3,5-Lutidine	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
49	120	3-Chloropyridine	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
50	73	1-Acetoxypropane	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
51	74	1-Acetoxybutane (Butylacetate)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
52	75	1-Acetoxypentane (Pentyl acetate)	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
53	57	1,1,1-Trifluorooctane	-2.3	-2.3	-2.3	-2.3	-2.3	-2.3	-2.3	-2.3
54	111	Fluorobenzene	-0.7	-0.7	-0.7	-0.7	-0.7	-0.7	-0.7	-0.7
55	112	Hexafluorobenzene	-2.9	-2.9	-2.9	-2.9	-2.9	-2.9	-2.9	-2.9
56	113	Trifluoromethylbenzene	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0
57	107	Dichloromethane	5.9	5.9	5.9	5.9	5.9	5.9	5.9	5.9
58	108	Trichloromethane	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6
59	109	Tetrachloromethane	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
60	59	1-Chlorobutane	-0.7	-0.7	-0.7	-0.7	-0.7	-0.7	-0.7	-0.7
61	60	1-Chloropentane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
62	61	1-Chlorohexane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
63	114	Chlorobenzene	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
64	62	1-Bromopropane	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5
65	63	1-Bromobutane	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3
66	64	1-Bromopentane	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2
67	115	Bromobenzene	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
68	88	1-Butanethiol	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
69	89	1-Pentanethiol	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
70	90	n-Hexanethiol	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
71	128	Thiophene	2.2	2.2	2.2	2.2	2.2	2.2	2.2	2.2
72	24	1-Hexene	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
73	47	Cyclohexene	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2
74	49	1,4 Cyclohexadiene	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
75	48	1,3 Cyclohexadiene	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
76	25	1-Heptene	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
77	26	1-Octene	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
78	27	1-Nonene	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
79	28	1-Decene	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
80	29	1-Pentyne	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
81	30	1-Hexyne	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1
82	35	2-Hexyne	-1.1	-1.1	-1.1	-1.1	-1.1	-1.1	-1.1	-1.1
83	36	3-Hexyne	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5
84	31	1-Heptyne	-0.8	-0.8	-0.8	-0.8	-0.8	-0.8	-0.8	-0.8
85	32	1-Octyne	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3
86	37	4-Octyne	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1
87	33	1-Nonyne	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2
88	34	1-Decyne	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2

ID [1]	ID [8]	Solutes	C78	РОН	TTF	MTF	PCL	тмо	PSH	PCN
89	50	Benzene	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
90	51	Toluene	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
91	52	Ethylbenzene	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
92	141	Naphthalene	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
93	142	Azulene	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
94	1	Pentane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
95	38	Cyclopentane	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4
96	11	2,2-Dimethylbutane	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1
97	12	2,3-Dimethylbutane	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
98	2	Hexane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
99	39	Cyclohexane	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1
100	13	2,2-Dimethylpentane	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4
101	14	2,3-Dimethylpentane	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
102	15	2,4-Dimethylpentane	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
103	20	2,2,3-Trimethylbutane	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
104	3	Heptane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
105	40	Cycloheptane	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
106	42	Methylcyclohexane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
107	17	2,3-Dimethylhexane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
108	18	2,4-Dimetylhexane	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
109	19	3,4-Dimetylhexane	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
110	21	2,2,4-Trimethylpentane	-0.7	-0.7	-0.7	-0.7	-0.7	-0.7	-0.7	-0.7
111	22	2,3,4-Trimethylpentane	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5
112	43	cis 1,2-Dimethylcyclohexane	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
113	44	trans 1,2 Dimethylcyclohexane	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
114	45	cis 1,4-Dimethylcyclohexane	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
115	46	trans 1,4 Dimethylcyclohexane	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
116	4	Octane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
117	41	Cyclooctane	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
118	5	Nonane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
119	6	Decane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
120	135	Cyclodecane	-1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
121	136	cis-Hydrindane	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
122	137	trans-Hydrindane	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
123	138	cis-Decalin	-0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
124	139	trans-Decalin	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
125	140	Adamantane	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
126	7	Undecane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
127	8	Dodecane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
128	9	Tridecane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
129	10	Tetradecane	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
130	121	Tetramethylsilane	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
131	122	Hexamethyldisilane	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
132	123	Hexamethyldisiloxane	1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.7
133	125	Tetramethylthin	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8

(Table 2) contd.....



Fig. (2). Diagram of the INPUT/OUTPUT of the Multiplicative Matrix Analysis (MMA), according to Laffort *et al.* [1] (see text).

In practice, the program starts with a classical multilinear regression, with R as independent variables and A as dependent variables. A first matrix B of solvent parameters is obtained, which in turn is considered as fixed. A first comparison between R and the product A * B is made. In a second step, a multi-linear regression is applied to B and R and a second value of matrix A of solute parameters is obtained. Further similar steps are used until two successive cycles do not provide an improvement in reconstruction of matrix R (it is experimentally observed that the system is convergent).

A revised version of the MMA programme in MATLAB language, previously published in [1], can be seen in Appendix A. Otherwise, an updated MMA applet is freely available in the web site: http://paul.laffort.free.fr

Solute Parameters

The MMA algorithm has been applied to the experimental R matrix of GLC data, as reported in Table SI-3 (127 x 11). Two input matrices A (127 x 5) have been tested: i) the solubility parameters according to [1], and ii) a randomized data matrix. The results are summarized in Table 3, in which δ , ω , ε , α and β respectively stand for the solute parameters of dispersion, orientation, induction /polarizability, acidity and basicity.

Table 3.Comparisons of Input and Output Solute
Parameters Using the MMA Algorithm Applied to
the Table SI-3, with Two Input Options (r and b
Respectively Stand for Correlation Coefficient and
for Multiplicative Coefficient)

Input	δ	ω	ε	α	β
(r) randomized	-0.03	0.04	0.01	0.18	0.07
(r) according to [1]	0.998	0.999	0.994	0.999	0.999
(b) according to [1]	0.98	1.00	1.01	1.00	1.00

Obviously, the input matrix according to [1] and the corresponding output matrix are very similar. Therefore, the solute parameters published in [1] are validated without any change, the minimal differences being in the same order as

those due to the experimental uncertainty in this type of measurements. Is also consequently validated the pooling experimental data set for 456 solutes, published in [3] and containing the previous data set of [1]. However, in order to be available for possible further studies, the solvation parameters for 127 solutes newly obtained are reported in Table **SI-4** of Supplementary Information.

Solvent Parameters of Stationary Phases

The things are slightly less simple. First of all, an erratum must be mentioned in Table 2 of [1]: the names of Cinf and C78 have been inverted in the two first lines of this table. This means that the reference (exhibiting the lowest values of solvent parameters) is C78, not Cinf. We observe the same corrected trend in the present study, as it can be seen in Table 4, where D, W, E, A and B respectively stand for the solvent parameters associated to the solute parameters δ , ω , ε , α and β .

In this Table 4, the values highlighted underline the most appropriate stationary phases for establishing solute solvation parameters with almost the same accuracy as with 10 or 11 phases, as demonstrated in [1]. Because these phases have been synthesized in limited quantity and are, therefore, not commercially available, the above results should be adapted and tested using more common phases. It clearly appears, as in [1], that the best set of commercial phases for further determinations of solute parameters from GLC, should include two non-polar phases of very different molecular weight and three polar phases: a strongly fluorinated, a classical polyether and an alcoholic (as, for example, diglycerol).

The range of correlations between the five solvent parameters reported in this Table 4 and those of Table 2 in [1] is 0.975-0.997. This fact can be interpreted as an improvement in the definition of the solvent parameters, on the contrary to that observed for solutes.

Table 4.Solvent Solvation Parameters of the 11 Stationary
Phases Under Study, Using the MMA Algorithm
Applied to the Experimental Matrix SI-3 (See Text)

Phase	D	W	Е	Α	В
C78	204.0	66.2	283.2	4.5	52.2
Cinf	204.2	68.3	306.6	4.7	54.4
РОН	204.1	86.8	291.5	27.8	149.6
TTF	204.2	141.4	284.8	9.9	112.0
MTF	204.1	88.1	283.0	8.3	71.2
PCl	204.2	85.0	290.1	10.3	67.5
PBr	204.2	83.9	291.5	10.4	69.6
TMO	204.2	122.7	305.8	68.8	97.3
PSH	204.1	81.4	293.5	12.0	66.1
PCN	203.9	124.6	301.2	36.5	99.9
SOH	204.1	87.2	289.1	26.1	113.4

The rules for establishing the solvent parameters of any stationary phase, based on a limited number of solutes, have been given in eq. 21 of [1]. They are updated in Table 5, together with two alternative sets of regressive equations, in which octan, penton, hexon, nitrop, benz, azul, butol and lutid stand respectively for the retention indices (-100) of 1-octane, 2-pentanone, 2-hexanone, 1-nitropropane, benzene, azulene, 1-butanol and 3,4-lutidine. From the three statistical tests reported in Table 5, it clearly appears that four polar solutes (plus octane or constant), as applied in [1], are not completely sufficient and that five polar solutes are needed and sufficient. It is also observed in Table 5, that the so-called "present regression model" seems the most effective, particularly for E (the solvent parameter associated to the induction/solubility parameter of solutes ε). Therefore, the "present regression model" can be considered as a simple

and accurate method to classify GLC stationary phases. However, because more values might be already available in the sets of solutes recommanded in [1] and [16], these two alternative groups of equations might also be useful.

FURTHER EXPERIMENTAL MEASUREMENTS

In the perspective of further determinations of solute parameters entirely from GLC experimentation, let us consider again the basic equation according to Rohrschneider [17], which can be written, as we saw, in the following matricial form:

$$R = A^*B \tag{3}$$

in which R stands for a matrix of retention indices (RI)

 Table 5.
 Regressive Equations of the Solvation Parameters of 11 Stationary Phases, as they are Expressed in Table 4, vs the Retention Indices of Limited Number of Solutes (*). According to these Three Tests, the So-Called "Present Regression Model" Seems the Most Suitable

Present Regression Model						r	SEE	F
D = +0.2916 octan								
W = - 0.9339 octan	+ 1.6855 hexon 2650		+ 0.3121 azul 137		- 0.8196 lutid 869	0.999	1.0	2077
E = - 0.5808 octan		- 0.2007 nitrop 334	+ 0.7499 azul 1438	+ 0.0740 butol 32	- 0.1724 lutid 168	0.999	0.6	523
A = -0.0866 octan				+ 0.9385 butol 2733	- 0.4494 lutid 528	0.999	0.9	2091
B = -0.2500 octan	- 0.3479 hexon 53		- 0.7981 azul 315	+ 0.1763 butol 31	+ 1.5123 lutid 1166	0.999	1.4	1137
Regression Equations with	the Solutes Selected	by McReynolds,	, 1970 [16]			r	SEE	F
D = +0.2916 octan								
W = - 0.6603 octan	+ 0.7744 penton 145	+0.4802 nitrop 123	+ 0.5340 benz 37	Ž	-0.7343 pyr 385	1.000	0.6	4481
E = +0.2897 octan	-2.4663 penton 177	+1.1425 nitrop 150		-0.3136 butol 27	+1.4417 pyr 105	0.988	1.7	63
A = -0.0951 octan				+ 1.0460 butol 661	-0.7118 pyr 154	0.997	1.7	628
B = +0.1269 octan			-2.7707 benz 143		+2.2556 pyr 576	0.996	3.2	443
Regression Equations with	the Solutes Selected	by Laffort <i>et al.</i> ,	, 2005 [1]			r	SEE	F
D = +0.2916 octan								
W = - 0.4576 octan		+0.8927 nitrop 688			-0.2346 pyr 40	0.997	2.2	626
E = +0.6122 octan	-2.6942 hexon 155	+1.2286 nitrop 134		-0.3900 butol 32	+1.6946 pyr 101	0.987	1.8	55
A = -0.0951 octan				+ 1.0460 butol 661	-0.7118 pyr 154	0.997	1.7	628
B = -1.5791 octan	+3.1594 hexon 547	-1.9669 nitrop 411		+ 0.5348 butol 135		0.997	2.7	427

(*) The numbers placed under the regressive terms stand for their partial F ratios. Only the terms with partial F ratios > 25 have been taken into consideration. The values of three statistical tests for each regressive equation have been reported: Pearson correlation coefficient (r), standard error of estimate (SEE) and global F ratio (F). See more details in text.

Solvation Parameters. Part 4

diminished of 100 (the RI value for methane), and A and B respectively for parameters matrices of solutes and stationary phases. Three steps can be considered: fixed A values, calibration and selection of B phases, and R experimental measurements.

Fixed Values of Matrix A

The reference solute parameters are those reported in Table **SI-4** (Supplementary Information). Those needed for applying the model on the top of Table **5**, are reproduced in Table **6**.

Table 6.	Fixed Solvat	ion	Parameters	for 5	Reference Solutes
	According	to	Table	SI-4	(Supplementary
	Information)				

Compounds	δ	ω	ε	α	β
Octane	3.430	0.000	0.000	0.000	0.000
2-Hexanone	2.611	0.689	0.118	-0.080	0.355
1-Nitropropane	1.922	1.151	0.290	0.137	0.173
Azulene	3.663	0.369	1.563	-0.103	0.192
1-Butanol	1.825	0.154	0.304	0.999	0.382
3,4-Lutidine	2.884	0.337	0.819	-0.190	0.798

Calibration and Selection of B Phases

It should be underlined that a determination of solvation parameters solely from GLC experimentation was never satisfactory before the publication [1] and the corresponding thesis [18]. This observation includes our older attempts [15, 19], and those of other authors recently reviewed by Abraham *et al.* [20]. In all these latter cases, none of the selected stationary phases are proton donor enough, and deficiencies are often noted for the solvent parameter called E in the present study (associated with the polarizabilityinduction solute parameter ε). This last deficiency could be overcome, as we already saw, using two apolar phases of very different molecular weight (paraffins C67 and C103, as in the present study, appears suitable).

In order to derive with accuracy solvation parameters of solutes solely from GLC data, the next step needs to select and calibrate a suitable set of commercial phases. For that purpose it is suggested to apply the matricial equation 3, in which the matrix A is given by Table 6, and R results from a preliminary proceeding by trial and error measurements using these six solutes, until to get B values resembling to those of Table 4. The proton donor phase (probably alcoholic) should be strong enough, but not excessively in order to avoid chemical bonds. Any phase producing similar B values to those of the primary alcohol synthetized and applied by the Kováts group, as reported in Table 4, should be very probably suitable.

It is certainly not too much difficult to reach this goal using filled columns. With open tubular columns, the purpose is perhaps harder, concerning precisely the proton donor phase and the two apolar phases of very different molecular weight. This difficulty is suggested, for example, by the results of Poole *et al.* [21, 22].

Once the selection of stationary phases achieved, applicable to filled or open tubular columns, a set of five equations similar to that on the top of Table **5** can be easily established applying again the matricial equation 3. This new

set of equations could be applicable, not only to get new solute solvation parameters, but also to verify the stability of the system (obtention of constant values for the reference solutes).

Experimentation Strictly Speaking

The values for reference solutes A being fixed, and the way for calibrating and selecting new B phases being clarified, the requirements are fitted for new GLC experimentation providing additional precise solvation parameters of solutes.

DISCUSSION

Rectification of a Previous Hypothesis

In our last publication on solvation parameters [3], we established that our solute solvation parameter called δ , identified to the molar polarizability, can be obtained from the published parameters by Abraham and co-authors *via* two ways, one including the molar volume Vx according to the definition of Abraham and McGowan [23], and the other one including the partition coefficient air-hexadecane L16. We observed that for solutes in liquid state at room temperature the data obtained *via* the two ways are almost identical, whereas often it is not the case for gases and solids. We suggested a hypothesis of explanation concerning solids, by the possible determination of the refractive index in crystalline form. In fact, Abraham and co-authors never used those values of experimental refractive index for solids, but predicted from models¹.

Temperature Dependence

Kováts and Kresz [8] have published, in addition to updated values of retention indices (as reported in Table **SI-2** for 133 solutes on 9 stationary phases), three corrected thermodynamic constants related to the molal Henry coefficient (for the same set of solutions). One of these constants corresponds to the non temperature dependent contribution to the Henry coefficient (the partial molar enthalpy of transfer), whereas the two other constants correspond to contributions associated with the temperature (the partial molar entropy of transfer, and the molar heat capacity difference of solutes between the ideal solution and the ideal gas phase).

We have not tried to derive from these constants, the more or less temperature dependence of each of the solvation parameters of solutes. We have however compared the results of multilinear regression analysis of saturated vapor pressures of solutes at 130°C and 25°C as functions of the five solute solvation parameters. It clearly appears from this comparison that the proton donor parameter is more temperature dependent than the three other polar parameters (i.e.: orientation, induction-polarizability and proton acceptor).

Another aspect of the temperature experimentation is when the low volatility of solutes under study needs an higher owen temperature than that for which the solvent parameters of phases have been established (130°C in the present study). In that case, a second set of equations analogous to those of the top of Table **5** should be established, with the same or alternative phases.

Advantages and Disadvantages of the Various Methods of Solvation Parameters Determination

Because of the limited number of published GLC data using proton donor phases, as well as apolar phases of very different molecular weight, various authors, and principally Abraham and

¹ M.H. Abraham, University College, London, Pers. Com., 2008.

co-authors since 1990 [24], combine chromatographic experimental data (including sometimes reverse phase in HPLC) with watersolvent partition coefficients, refractive indices and molar volumes. The three additional sources of information are sometimes experimental and sometimes derived from the molecular structure.

A further determination of solute parameters entirely from GLC experimentation, presently suggested, could be a better alternative method, even using packed columns.

GENERAL CONCLUSION

The solute solvation parameters computed on the basis of the updated chromatographic data by Kováts and Kresz [8] for 127 compounds, are almost identical to those published in [1]. By contrast, we observe a slight improvement in the definition of solvent solvation parameters for the 11 involved stationary

phases. We have specified from these results a simple method

of stationary phases classification and a method for further determination of solute solvation parameters.

ACKNOWLEDGEMENTS

The authors are indebted to the Kováts group in Lausanne, for their experimental work (synthesis and chromatography) on which the reference [1] and the present study are based. Many thanks again to Ervin Kováts itself, and also to R. Cloux, A. Dallos, C. De Reyff, G. Défayes, J.C. Dutoit, K. Fóti, D.F. Fritz, T. Görner, G.A. Huber, K.S. Reddy, F. Riedo, G. Tarján, and more recently in Veszprém (Hungary), to R. Kresz, Z. Kulksár and A. Sizak.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's web site along with the published article.

APPENDIX A

MMA Programme in MATLAB Language

```
function [Ap,B,Error,Correlations] = mma(pathA,pathR,epsilon)
  A=load('-ascii',pathA);
        R=load('-ascii',pathR);
        [x,y,u,v]=solve(A,R,epsilon);
        Ap=x:
        B=y;
        Error=u;
        Correlations=v;
end;
function [Ap,B,Error,Correlations] = solve(A,R,epsilon)
        [nlr,ncr] = size(R); \% dimension matrix R
        [nla,nca] = size(A) \% dimension matrix A
        N = nlr*ncr; % number of elements of matrix R
        oldA=A:
        B = rand(nca,ncr); % initialize matrix B
        count=1;
                                          % initialize count
        test=1;
        B=pinv(A)*R;
        AB=A*B;
        A=R*pinv(B);
        SSD= sum(sum( (R - A*B) .^ 2)); %sum of squares of differences between AB and R
        Error(count) = sqrt(SSD / (N-1));
                                            %standard error
        while test > epsilon
                count=count+1;
                B=pinv(A)*R;
                AB=A*B:
                A=R*pinv(B);
                SSD= sum(sum( (R - A*B) ^2); %sum of squares of differences between AB and R
                Error(count) = sqrt(SSD / (N-1));
                                                     %standard error
                test=Error(count-1)-Error(count);
        end:
        Ap=A;
        Correlations=corrx(oldA,A);
end:
function [res] = corrx(M,N)
  [nlM.ncM] = size(M):
  d=sum(M.*N)-((sum(M).*sum(N))/nIM);
  n = sqrt((sum(N.*M)-((sum(M).*sum(M))/nlM)).*(sum(N.*N)-((sum(N).*sum(N))/nlM)));
  res=d./n;
end:
```

REFERENCES

- P. Laffort, F. Chauvin, A. Dallos, P. Callegari, and D. Valentin, "Solvation parameters. Part 1: mutual improvements of several approaches and determination of two first sets of optimized values", *J. Chromatogr. A*, vol. 1100, pp. 90-107, 2005.
- [2] P. Laffort, and P. Héricourt, "Solvation parameters.2: a simplified molecular topology to generate easily optimized values", J. Chem. Inf. Model, vol.46, pp. 1723-1734, 2006.
- [3] P. Laffort, and P. Héricourt, "Solvation parameters of solutes: an attempt of improvements in characterization and determination", *TOAINFOJ*, vol. 2, pp. 22-31, 2008.
- [4] K. S. Reddy, J. C. Dutoit, and E. sz. Kováts, "Pair-wise interactions by gas chromatography. Part I: interaction free enthalpies of solutes with non associated primary alcohol groups", *J Chromatogr.*, vol. 609, pp. 229-259, 1992.
- [5] K. S. Reddy, R. Cloux, and E. sz. Kováts, "Pair-wise interactions by gas chromatography. Part IV: interaction free enthalpies of solutes with trifluoromethyl-substituted alkanes", *J Chromatogr.*, vol. 673, pp.181-209, 1994.
- [6] G. Défayes, K. S. Reddy, A. Dallos, and E. sz. Kováts, "Pair-wise interactions by gas chromatography. Part V: interaction free enthalpies of solutes with primary chloro- and bromo-alkanes", J. *Chromatogr.*, vol. 699, pp. 131-154, 1995.
- [7] K. S. Reddy, R. Cloux, and E. sz. Kováts, "Pair-wise interactions by gas chromatography. Part VI: interaction free enthalpies of solutes with primary methoxyalkane, cyanoalkane, and alkanethiol groups", J. Chromatogr., vol. 704, pp. 387-436, 1995.
- [8] E. sz. Kováts, and R. Kresz, "Wrong gas/liquid partition data by gas chromatography", J. Chromatogr. A, vol. 1113, pp. 206-219, 2006.
- [9] R. Cloux, G. Défayes, K. Fóti, J. C. Dutoit, and E. sz. Kováts, "Pair-wise interactions by gas chromatography. Part III: synthesis of isosteric stationary phases for gas-chromatography", *Synthesis*, pp. 909-919, 1993.
- [10] F. Riedo, D. Fritz, G. Tarján, and E. sz. Kováts, "A tailor-made C87 hydrocarbon as a possible non-polar standard stationary phase for gas chromatography", *J. Chromatogr.*, vol. 126, pp. 63-83, 1976.
- [11] G. Défayes, D. F. Fritz, T. Görner, G. Huber, C. De Reyff, and E. sz. Kováts, "Organic solutes in paraffin solvents: influence of the size of the solvent molecule on solution data", *J Chromatogr.*, vol. 500, pp. 139-184, 1990.

Received: March 27, 2009

Revised: May 4, 2009

Accepted: May 7, 2009

© Laffort and Héricourt; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/ 3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.

- [12] K. S. Reddy, and E. sz. Kováts, "Dependence of solution data at ideal dilution on solvent molecular size", *Chromatogr.*, vol. 30, pp. 493-499, 1990.
- [13] D. E. Martire, "Dependence of solute retention parameters on stationary phase molecular weight in gas-liquid chromatography", *Anal. Chem.*, vol. 46, pp. 626-633, 1974.
- [14] A. Dallos, A. Sisak, Z. Kulcsár, and E sz. Kováts, "Pair-wise interactions by gas chromatography. VII: interaction free enthalpies of solutes with secondary alcohol groups". J. Chromatogr. A, vol. 904, pp 211-242, 2000.
- [15] P. Laffort, and F. Patte, "The solubility factors in gas-liquid chromatography: comparison between two approaches and application to some biological studies". J. Chromatogr., vol. 126, pp. 625-639, 1976.
- [16] W. O. McReynolds, "Characterization of some liquid phases", J. Chromatogr. Sci., vol. 8, pp. 685-691, 1970.
- [17] L. Rohrschneider, "Eine methode zur charakterisierung von gaschromatographischen Trennflüssigkeiten", J. Chromatogr., vol. 22, pp. 6-22, 1966.
- [18] F. Chauvin, "Improvement in the definition and determination of the solvation parameters - Application to olfaction (in French)", Thesis University of Bourgogne, Dijon, France, 1998; p. 149.
- [19] F. Patte, M. Etcheto, and P. Laffort, "Solubility factors for 240 solutes and 207 stationary phases in Gas-Liquid Chromatography", *Anal. Chem.*, vol. 54, pp. 2239-2247, 1982.
- [20] M. H. Abraham, A. Ibrahim, and A. M. Zissimos, "Determination of sets of solute descriptors from chromatographic measurements", *J. Chromatogr. A*, vol. 1037, pp. 29-47, 2004.
- [21] C. F. Poole, Q. Li, W. Kiridena, and W. W. Koziol, "Selectivity assessment of popular stationary phases for open-tubular column gas chromatography. J. Chromatogr. A, vol. 912, pp. 107-117, 2001.
- [22] C. F. Poole, and S. K. Poole, "Separation characteristics of wallcoated open-tubular columns for gas chromato-graphy", J. Chromatogr. A, vol. 1184, pp. 254-280, 2008.
- [23] M. H. Abraham, and J. C. McGowan, "The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography", *Chromatographia*, vol. 23, pp. 243-246, 1987.
- [24] M. H. Abraham, G. Whiting, R. M. Doherty, and W. H. Shuely, "Hydrogen bonding. Part 13: a new method for the characterisation of GLC stationary phases-The Laffort data set", J. Chem. Soc. Perkin Trans. I., vol. 2, pp. 1451-1460, 1990.