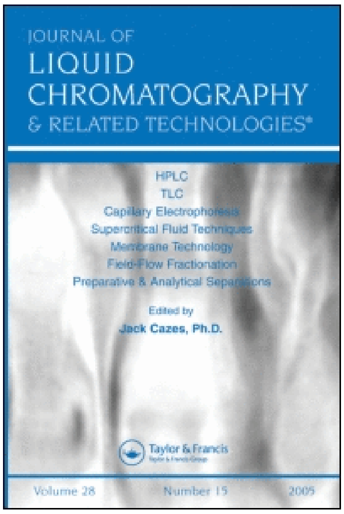


Journal of Liquid Chromatography & Related Technologies
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

Application of Solvatochromic Parameters to Selectivity Tuning in Chromatography for Aromatic Solutes

F. Hellal[a]; R. Phan-Tan-Luu[a]; A. M. Siouffi[a]
[a] Faculté des Sciences de Saint-Jerôme, Université d'Aix-Marseille III, Marseille, Cedex, France

To cite this Article Hellal, F. , Phan-Tan-Luu, R. and Siouffi, A. M.(1994) 'Application of Solvatochromic Parameters to Selectivity Tuning in Chromatography for Aromatic Solutes', Journal of Liquid Chromatography & Related Technologies, 17: 13, 2845 — 2869

To link to this Article: DOI: 10.1080/10826079408013504

URL: http://dx.doi.org/10.1080/10826079408013504

JOURNAL OF LIQUID CHROMATOGRAPHY, 17(13), 2845-2869 (1994)

APPLICATION OF SOLVATOCHROMIC PARAMETERS TO SELECTIVITY TUNING IN CHROMATOGRAPHY FOR AROMATIC SOLUTES

F. HELLAL, R. PHAN-TAN-LUU, AND A. M. SIOUFFI
Faculté des Sciences de Saint-Jerôme
Université d'Aix-Marseille III
F-13397 Marseille Cedex 20, France

ABSTRACT

Solvatochromic parameters are the best descriptors of solute retention in RPLC. Accurate prediction of capacity factors for 26 solutes is obtained with binary mixtures of organic modifier/water in the vol/vol range 10-100%. Example of application is given.

INTRODUCTION

The exact mechanism of retention in reversed-phase liquid chromatography (RPLC) has been the subject of much controversy and debate. Ideally, knowledge of this mechanism would allow a priori prediction of retention times regardless of the column being used, as well as computer-based optimization of a given separation. The difficulty in elucidating the mechanism of retention lies in the numerous interactions that a solute may undergo in both stationary and mobile phases.

Retention is closely related to solute shape and ability to form weak complexes. In order to describe molecular properties of a solute, it is necessary that those parameters should exhibit additive and constitutive properties which can be obtained not only from experimental measurements but also from theoretical calculations. The descriptors which satisfy the above requirements were selected from many published sources and are as follows:

- **LogP**, the logarithm of the partition coefficient in 1-octanol-water, which is a measure of the hydrophobicity of the molecule [1,2]. LogP is determined from shake flask or slow stirring[3]. Some data are from RPLC and cannot be considered.

- **RI**, retention index of Burr and Smith[4-9]. The basis of the prediction system proposed in their study is that the retention index of an analyte in a selected eluent can be calculated by the summation of the retention index of a parent compoud(PI), substituent index values (SI) for each substituent plus terms required to describe interactions between substituents (interaction indices II, i.e., hydrogen bonding, steric and electronic interaction). The retention index of a compound can then be determined as

$$\mathbf{RI = PI + SI_{R} + \sum SI_{Ar\text{-}X} + \sum SI_{R\text{-}X} + \sum II_{YZ}}$$

PI : Retention index value of a parent compound

SI_R : The retention index contribution from saturated aliphatic carbons

$\Sigma SI_{AR\text{-}X}$: Substituent index values for substituents on an aromatic ring

$\Sigma SI_{R\text{-}X}$: Substituent index values for subtituents on saturated aliphatic carbons

ΣII_{YZ} : Interaction index values between substituents to account for H-bonding, and electronic effects.

- Van der Waals volume, V_w, and Van der Waals surface area, A_W, calculated from the Van der Waals radii of the atoms of which the molecule is composed[10].

- Molecular connectivity index, χ, which measures the topological size of the molecule and its degree of branching[11].

- Correlation factor, F, calculated as (number of double bonds) + (number of primary and secondary carbon atoms) - 0.5 for a non-aromatic ring[12,13].

- Solvatochromic parameters proposed by Kamlet et al[14-16] which are :

V_i the cavity term, is a measure of solute volume and may be V, the liquid molar volume, taken as the solute molecular weight divided by its liquid density at 20°C, or Vi, the intrinsic (Van der Waals) molar volume which can be either computer calculated or estimated by simple additivity methods like that of McGowan[17].

. π^* solvatochromic parameter is a measure of solute dipolarity/polarizability.

. β and α solvatochromic parameters are a measure of solute basicity (hydrogen bond acceptor) and acidity (hydrogen bond donor), respectively.

Recently the linear solvation energy relationship (LSER) based on the Kamlet Taft multiparameter scale has been successfuly used to model retention in RPHPLC[18,19].

In our aproach, described in previous works (20-22), we consider the phenomenon as a dark box (Fig. 1) and we try to correlate the relevant factors to the answer but no theoretical retention model is considered as the right one.

The purpose of this paper is to select those factors which can permit the determination of accurate answers.

CHOICE OF THE DESCRIPTORS

To avoid duplication of information that is provided by the variables, we must perform relevant choice of the solute descriptors from those enumerated above.

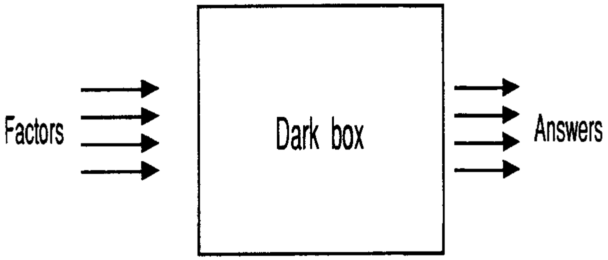


FIGURE 1. Adopted approach

logP and logk' are varying linearly with Vi/100, π^*, β and α according to:

$$\log P = b_0 + b_1 Vi/100 + b_2 \pi^* + b_3 \beta + b_4 \alpha^{(16)} \qquad (1)$$

$$\log k' = b'_0 + b'_1 Vi/100 + b'_2 \pi^* + b'_3 \beta + b'_4 \alpha^{(19)} \qquad (2)$$

with a fixed percentage of organic modifier which implies a similar equation for RI. Since $\log k' = a_0 + a_1 \, RI^{(9)}$ (3)

we can write : $RI = b''_0 + b''_1 Vi/100 + b''_2 \pi^* + b''_3 \beta + b''_4 \alpha$ (4)

Subsequently we will discard logP and RI from the solute descriptors.

From a literature survey we could retrieve some relevant data which are compiled in Table 1. It must be pointed out that solutes are all aromatic compounds.

From Table 1 we can construct the following correlation matrix(24) (Table 2).

The scrutinizing of the off-diagonal elements of the correlation matrix, reveals the independence of the four descriptors (Vi/100, π^*, β and α) and the dependence between Vi/100 and the variables (χ , Vw, Aw and F). From these conclusions we can select the four variables Vi/100, π^*, β and α as solute descriptors only (Fig 2).

RESULTS AND DISCUSSION

From many experimental data it has been demonstrated(25-28) that lnk' varies quadratically with the percent of organic modifier :

$$\ln k' = A \Phi^2 + B \Phi + C \qquad (5)$$

Φ is the percent of organic modifier.

At a constant percentage of organic modifier many researchers found a linear relationship between lnk' and the four factors Vi/100, π^*, β, α (19,29,30) :

$$\ln k' = b_0 + b_1 (Vi/100) + b_2 \pi^* + b_3 \beta + b_4 \alpha. \qquad (6)$$

To take into account interactions between solute and mobile phase the general model can be obtained by forming the product of the two models proposed above:

$$\ln k' = (b_0 + b_1 (Vi/100) + b_2 \pi^* + b_3 \beta + b_4 \alpha) * (A \Phi^2 + B \Phi + C) \qquad (7)$$

which is equivalent to :

TABLE 1
Selected Solute Descriptors in RPLC from Literature Data.

	Solute descriptors							
Solute	**F**	**χ**	**V_w**	**A_w**	**$V_i/100$**	**π^***	**β**	**α**
Benzene	3	2.00	48.36	6.01	0.49	0.59	0.10	0.00
Toluene	4	2.41	59.51	7.45	0.59	0.55	0.11	0.00
Ethylbenzene	5	2.97	69.74	8.80	0.67	0.53	0.12	0.00
o-Xylene	5	2.83	70.66	8.89	0.67	0.51	0.12	0.00
m-Xylene	5	2.82	70.66	8.89	0.67	0.51	0.12	0.00
n-Propylbenzene	6	3.47	79.97	10.15	0.77	0.51	0.12	0.00
Isopropylbenzene	5	3.35	79.96	10.14	0.77	0.51	0.12	0.00
n-Butylbenzene	7	3.97	90.20	11.50	0.87	0.49	0.12	0.00
p-Cymene	6	3.77	91.11	11.58	0.87	0.47	0.13	0.00
1,2,4,5-Tetramethylbenzene	7	3.66	92.96	11.77	0.87	0.43	0.15	0.00
Naphtalene	5	3.41	73.96	8.42	0.75	0.70	0.15	0.00
Acenaphthene	6.5	4.45	87.88	9.54	0.92	0.66	0.17	0.00
Anthracene	7	4.81	99.56	10.84	1.02	0.80	0.20	0.00
Phenanthrene	7	4.82	99.56	10.84	1.02	0.80	0.20	0.00
Pyrene	8	5.56	109.04	11.26	1.16	0.90	0.25	0.00
Benz(a)anthracene	9	6.22	125.16	13.26	1.28	0.90	0.25	0.00
Aniline	3	2.20	56.38	7.07	0.56	0.73	0.50	0.26
N-Ethylaniline	5	3.22	77.82	9.79	0.76	0.82	0.47	0.17
Benzaldehyde	3	2.44	60.06	7.61	0.61	0.92	0.44	0.00
Benzonitrile	3	2.38	60.54	7.52	0.59	0.90	0.37	0.00
Nitrobenzene	3	2.45	62.64	7.88	0.63	1.01	0.30	0.00
Anisole	4	2.52	62.71	7.99	0.639	0.73	0.32	0.00
Acetophenone	4	2.87	71.21	9.05	0.69	0.90	0.49	0.04
Methyl benzoate	4	2.98	76.73	9.83	0.74	0.75	0.39	0.00
m-Chloroaniline	3	2.71	65.86	8.20	0.65	0.78	0.40	0.31

Data of χ , Vw, Aw and F are taken from ref 23 and data of Vi/100, π^*, β and α are taken from ref 16.

TABLE 2
Correlation Matrix

	F	χ	V_W	A_W	$V_i/100$	π^*	β	α
F	1.000							
χ	0.937	1.000						
V_W	0.950	0.975	1.000					
A_W	0.904	0.859	0.947	1.000				
$V_i/100$	**0.934**	**0.994**	**0.989**	**0.891**	**1.000**			
π^*	-0.167	0.127	0.048	-0.127	0.141	1.000		
β	-0.441	-0.242	-0.235	-0.254	-0.201	0.723	1.000	
α	-0.357	-0.254	-0.257	-0.255	-0.246	-0.182	0.585	1.000

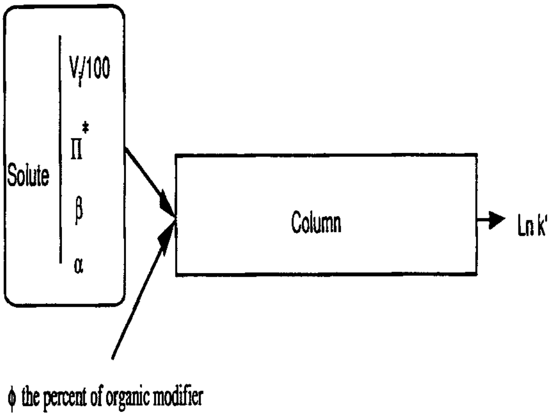


FIGURE 2. Selected factors

TABLE 3
Solvatochromic Parameters Values of 15 Solutes Taken from ref. 16

	Solute	$V_I/100$	π^*	β	α	Φ
1	Acetophenone	0.690	0.90	0.49	0.04	10
2	Aniline	0.562	0.73	0.50	0.26	10
3	Anisole	0.639	0.73	0.32	0.00	10
4	Anthracene	1.015	0.80	0.20	0.00	10
5	Benzaldehyde	0.606	0.92	0.44	0.00	10
6	Benzene	0.491	0.59	0.10	0.00	10
7	Benzonitrile	0.590	0.90	0.37	0.00	10
8	Benzyl alcohol	0.634	0.99	0.52	0.39	10
9	Chlorobenzene	0.581	0.71	0.07	0.00	10
10	Ethylbenzene	0.668	0.53	0.12	0.00	10
11	Nitrobenzene	0.631	1.01	0.30	0.00	10
12	Phenol	0.536	0.72	0.33	0.61	10
13	2-Phenylethanol	0.732	0.97	0.55	0.33	10
14	3-Phenylpropanol	0.830	0.95	0.55	0.33	10
15	Toluene	0.592	0.55	0.11	0.00	10
16	Acetophenone	0.690	0.90	0.49	0.04	32.5
17	Aniline	0.562	0.73	0.50	0.26	32.5
18	Anisole	0.639	0.73	0.32	0.00	32.5
19	Anthracene	1.015	0.80	0.20	0.00	32.5
20	Benzaldehyde	0.606	0.92	0.44	0.00	32.5
21	Benzene	0.491	0.59	0.10	0.00	32.5
22	Benzonitrile	0.590	0.90	0.37	0.00	32.5
23	Benzyl alcohol	0.634	0.99	0.52	0.39	32.5
24	Chlorobenzene	0.581	0.71	0.07	0.00	32.5
25	Ethylbenzene	0.668	0.53	0.12	0.00	32.5
26	Nitrobenzene	0.631	1.01	0.30	0.00	32.5
27	Phenol	0.536	0.72	0.33	0.61	32.5
28	2-Phenylethanol	0.732	0.97	0.55	0.33	32.5
29	3-Phenylpropanol	0.830	0.95	0.55	0.33	32.5
30	Toluene	0.592	0.55	0.11	0.00	32.5
31	Acetophenone	0.690	0.90	0.49	0.04	55
32	Aniline	0.562	0.73	0.50	0.26	55
33	Anisole	0.639	0.73	0.32	0.00	55
34	Anthracene	1.015	0.80	0.20	0.00	55
35	Benzaldehyde	0.606	0.92	0.44	0.00	55
36	Benzene	0.491	0.59	0.10	0.00	55
37	Benzonitrile	0.590	0.90	0.37	0.00	55
38	Benzyl alcohol	0.634	0.99	0.52	0.39	55
39	Chlorobenzene	0.581	0.71	0.07	0.00	55
40	Ethylbenzene	0.668	0.53	0.12	0.00	55
41	Nitrobenzene	0.631	1.01	0.30	0.00	55
42	Phenol	0.536	0.72	0.33	0.61	55
43	2-Phenylethanol	0.732	0.97	0.55	0.33	55
44	3-Phenylpropanol	0.830	0.95	0.55	0.33	55
45	Toluene	0.592	0.55	0.11	0.00	55

(continued)

TABLE 3 (Continued)
Solvatochromic Parameters Values of 15 Solutes Taken from ref. 16

	Solute	V_i /100	π*	β	α	Φ
46	Acetophenone	0.690	0.90	0.49	0.04	77.5
47	Aniline	0.562	0.73	0.50	0.26	77.5
48	Anisole	0.639	0.73	0.32	0.00	77.5
49	Anthracene	1.015	0.80	0.20	0.00	77.5
50	Benzaldehyde	0.606	0.92	0.44	0.00	77.5
51	Benzene	0.491	0.59	0.10	0.00	77.5
52	Benzonitrile	0.590	0.90	0.37	0.00	77.5
53	Benzyl alcohol	0.634	0.99	0.52	0.39	77.5
54	Chlorobenzene	0.581	0.71	0.07	0.00	77.5
55	Ethylbenzene	0.668	0.53	0.12	0.00	77.5
56	Nitrobenzene	0.631	1.01	0.30	0.00	77.5
57	Phenol	0.536	0.72	0.33	0.61	77.5
58	2-Phenylethanol	0.732	0.97	0.55	0.33	77.5
59	3-Phenylpropanol	0.830	0.95	0.55	0.33	77.5
60	Toluene	0.592	0.55	0.11	0.00	77.5
61	Acetophenone	0.690	0.90	0.49	0.04	100
62	Aniline	0.562	0.73	0.50	0.26	100
63	Anisole	0.639	0.73	0.32	0.00	100
64	Anthracene	1.015	0.80	0.20	0.00	100
65	Benzaldehyde	0.606	0.92	0.44	0.00	100
66	Benzene	0.491	0.59	0.10	0.00	100
67	Benzonitrile	0.590	0.90	0.37	0.00	100
68	Benzyl alcohol	0.634	0.99	0.52	0.39	100
69	Chlorobenzene	0.581	0.71	0.07	0.00	100
70	Ethylbenzene	0.668	0.53	0.12	0.00	100
71	Nitrobenzene	0.631	1.01	0.30	0.00	100
72	Phenol	0.536	0.72	0.33	0.61	100
73	2-Phenylethanol	0.732	0.97	0.55	0.33	100
74	3-Phenylpropanol	0.830	0.95	0.55	0.33	100
75	Toluene	0.592	0.55	0.11	0.00	100

$$\begin{aligned} \ln k' = {} & b_0' + b_1'\left(\frac{V_i}{100}\right) + b_2'\pi^* + b_3'\beta + b_4'\alpha \\ & + \left[b_0'' + b_1''\left(\frac{V_i}{100}\right) + b_2''\pi^* + b_3''\beta + b_4''\alpha\right]\Phi \\ & + \left[b_0''' + b_1'''\left(\frac{V_i}{100}\right) + b_2'''\pi^* + b_3'''\beta + b_4'''\alpha\right]\Phi^2. \end{aligned} \tag{8}$$

From the data of Table 3 we have the following experimental domain :

TABLE 4
Experimental Domain

Natural variables	Center	Range of variation
$V_i/100$	0.753	0.262
π^*	0.770	0.240
β	0.310	0.240
α	0.305	0.305
Φ	0.550	0.450

$$\text{Center} = \frac{\text{Highest natural variable limit} + \text{Lowest natural variable limit}}{2}$$

$$\text{Range of variation} = \frac{\text{Highest natural variable limit} - \text{Lowest natural variable limit}}{2}$$

It must be pointed out that the lowest limit of Φ is 10%. Beyond this value retention is by far too high and our purpose is not to determine $\log K_w$ values. $V_i/100$, π^*, β, α and Φ are the natural variables which can be transformed into coded variables X_1, X_2, X_3, X_4, X_5 respectively by the relation:

$$X_i = \frac{\text{natural variable - center}}{\text{range of variation, } r_i}$$

which yield the following relationship of lnk' versus the five parameters:

$$\begin{aligned} \ln k' = {} & b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 \\ & + [\,b_5 + b_{15}X_1 + b_{25}X_2 + b_{35}X_3 + b_{45}X_4]\;X_5 \\ & + [\,b_{55} + b_{155}X_1 + b_{255}X_2 + b_{355}X_3 + b_{455}X_4]\;X_5^2 \end{aligned}$$

TABLE 5
Matrix with coded variables

N° of experiment	Coded Variables					Answer values taken from ref. 26		
	X1	X2	X3	X4	X5	lnk' MeOH	lnk' MeCN	lnk' THF
1	-0.24	0.54	0.75	-0.87	-1	4.55	3.29	2.71
2	-0.73	-0.17	0.79	-0.15	-1	2.61	2.40	2.36
3	-0.44	-0.17	0.04	-1	-1	7.97	4.21	4.25
4	1	0.13	-0.46	-1	-1	11.58	6.51	6.56
5	-0.56	0.63	0.54	-1	-1	3.53	3.01	2.56
6	-1	-0.75	-0.87	-1	-1	4.50	4.16	4.38
7	-0.62	0.54	0.25	-1	-1	3.95	3.36	2.96
8	-0.45	0.92	0.87	0.28	-1	3.00	2.27	1.87
9	-0.66	-0.25	-1	-1	-1	7.28	5.33	5.48
10	-0.32	-1	-0.79	-1	-1	7.55	5.71	5.53
11	-0.47	1	-0.04	-1	-1	4.23	3.82	4.11
12	-0.83	-0.21	0.08	1	-1	2.82	2.52	2.94
13	-0.08	0.83	1	0.08	-1	3.90	3.06	2.51
14	0.29	0.75	1	0.08	-1	5.10	3.67	3.52
15	-0.61	-0.92	-0.83	-1	-1	6.80	5.19	5.28
16	-0.24	0.54	0.75	-0.87	-0.5	2.65	1.63	0.99
17	-0.73	-0.17	0.79	-0.15	-0.5	1.42	0.98	1.12
18	-0.44	-0.17	0.04	-1	-0.5	4.57	2.53	2.00
19	1	0.13	-0.46	-1	-0.5	8.75	4.61	3.36
20	-0.56	0.63	0.54	-1	-0.5	2.14	1.52	0.99
21	-1	-0.75	-0.87	-1	-0.5	3.19	2.48	2.21
22	-0.62	0.54	0.25	-1	-0.5	2.26	1.73	1.21
23	-0.45	0.92	0.87	0.28	-0.5	1.61	0.67	0.38
24	-0.66	-0.25	-1	-1	-0.5	4.82	3.24	2.73
25	-0.32	-1	-0.79	-1	-0.5	5.32	3.68	3.01
26	-0.47	1	-0.04	-1	-0.5	2.75	2.11	1.82
27	-0.83	-0.21	0.08	1	-0.5	1.52	0.91	1.16
28	-0.08	0.83	1	0.08	-0.5	2.24	1.08	0.66
29	0.29	0.75	1	0.08	-0.5	3.11	1.57	1.18
30	-0.61	-0.92	-0.83	-1	-0.5	4.60	3.21	2.75
31	-0.24	0.54	0.75	-0.87	0	1.13	0.40	-0.28
32	-0.73	-0.17	0.79	-0.15	0	0.41	-0.03	0.01
33	-0.44	-0.17	0.04	-1	0	2.05	1.21	0.32
34	1	0.13	-0.46	-1	0	5.91	2.95	0.96
35	-0.56	0.63	0.54	-1	0	0.94	0.37	-0.21
36	-1	-0.75	-0.87	-1	0	1.91	1.13	0.54
37	-0.62	0.54	0.25	-1	0	0.89	0.47	-0.13
38	-0.45	0.92	0.87	0.28	0	0.48	-0.39	-0.67
39	-0.66	-0.25	-1	-1	0	2.72	1.58	0.69

TABLE 5 (Continued)
Matrix with coded variables

N° of experiment	Coded Variables					Answer values taken from ref. 26		
	X1	X2	X3	X4	X5	lnk' MeOH	lnk' MeCN	lnk' THF
40	-0.32	-1	-0.79	-1	0	3.28	2.00	1.04
41	-0.47	1	-0.04	-1	0	1.41	0.75	0.14
42	-0.83	-0.21	0.08	1	0	0.43	-0.21	-0.19
43	-0.08	0.83	1	0.08	0	0.87	-0.22	-0.63
44	0.29	0.75	1	0.08	0	1.44	0.14	-0.43
45	-0.61	-0.92	-0.83	-1	0	2.68	1.61	0.82
46	-0.24	0.54	0.75	-0.87	0.5	-0.01	-0.40	-1.09
47	-0.73	-0.17	0.79	-0.15	0.5	-0.41	-0.63	-0.97
48	-0.44	-0.17	0.04	-1	0.5	0.39	0.25	-0.80
49	1	0.13	-0.46	-1	0.5	3.08	1.50	-0.64
50	-0.56	0.63	0.54	-1	0.5	-0.07	-0.42	-1.04
51	-1	-0.75	-0.87	-1	0.5	0.66	0.10	-0.62
52	-0.62	0.54	0.25	-1	0.5	-0.18	-0.41	-1.05
53	-0.45	0.92	0.87	0.28	0.5	-0.40	-0.91	-1.31
54	-0.66	-0.25	-1	-1	0.5	0.98	0.35	-0.65
55	-0.32	-1	-0.79	-1	0.5	1.42	0.68	-0.40
56	-0.47	1	-0.04	-1	0.5	0.22	-0.25	-0.94
57	-0.83	-0.21	0.08	1	0.5	-0.44	-0.84	-1.11
58	-0.08	0.83	1	0.08	0.5	-0.20	-0.83	-1.34
59	0.29	0.75	1	0.08	0.5	0.09	-0.61	-1.30
60	-0.61	-0.92	-0.83	-1	0.5	1.04	0.40	-0.49
61	-0.24	0.54	0.75	-0.87	1	-0.77	-0.77	-1.46
62	-0.73	-0.17	0.79	-0.15	1	-1.05	-0.83	-1.82
63	-0.44	-0.17	0.04	-1	1	-0.41	-0.33	-1.35
64	1	0.13	-0.46	-1	1	0.24	0.28	-1.43
65	-0.56	0.63	0.54	-1	1	-0.89	-0.86	-1.50
66	-1	-0.75	-0.87	-1	1	-0.56	-0.60	-1.27
67	-0.62	0.54	0.25	-1	1	-0.95	-0.91	-1.56
68	-0.45	0.92	0.87	0.28	1	-1.01	-0.88	-1.52
69	-0.66	-0.25	-1	-1	1	-0.40	-0.44	-1.28
70	-0.32	-1	-0.79	-1	1	-0.24	-0.30	-1.28
71	-0.47	1	-0.04	-1	1	-0.82	-0.90	-1.42
72	-0.83	-0.21	0.08	1	1	-1.09	-0.99	-1.61
73	-0.08	0.83	1	0.08	1	-0.97	-0.75	-1.48
74	0.29	0.75	1	0.08	1	-0.94	-0.68	-1.45
75	-0.61	-0.92	-0.83	-1	1	-0.33	-0.43	-1.20

According to the characteristics of the chosen matrix (see appendix n°1) this matrix cannot permit the calculation of the model coefficients with inflation factors greater than $7^{(31\text{-}34)}$ and the term b_{355} should be discarded:

$$\begin{aligned} \ln k' = {} & b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 \\ & + [\, b_5 + b_{15}X_1 + b_{25}X_2 + b_{35}X_3 + b_{45}X_4]\ X_5 \\ & + [\, b_{55} + b_{155}X_1 + b_{255}X_2 + b_{455}X_4]\ X_5^2 \end{aligned} \qquad (10)$$

Since methanol, acetonitrile and tetrahydrofran represent the usual organic modifiers we calculated the relationship between lnk' and the 5 factors for these 3 solvents.

Equations are as follows :

- in methanol-water (see appendix n°2):

$$\begin{aligned} \ln k'_M = {} & 2.582 + 2.255\,X_1 - 0.478\,X_2 - 0.994\,X_3 - 0.342\,X_4 \\ & + [\, -3.538 - 1.651\,X_1 + 0.359\,X_2 + 0.591\,X_3 + 0.305\,X_4]\ X_5 \\ & + [0.453\ - 0.125\,X_1 + 0.067\,X_2 - 0.049\,X_4]\ X_5^2 \end{aligned} \qquad (11)$$

According to the results of residual plot and the normal plot (see appendix n°2) we observe high deviation between calculated and observed anisol capacity factor at methanol-water (10/90) vol/vol.

$$\ln k'_{exp} - \ln k'_{cal} = 7.96 - 5.68 = 2.28. \qquad (12)$$

In fact, this value can be discarded since we do not work with k'= exp(7.96) = 2864 because the analysis time is beyond usual range, and this experiment can be withdrawn from the matrix without introducing significant changes to the coefficient values of the postulated model (see appendix n°3).

- in acetonitrile-water (see appendix n°4):

$$\begin{aligned} \ln k'_A = {} & 1.01 + 1.05\,X_1 - 0.34\,X_2 - 0.67\,X_3 - 0.47\,X_4 \\ & + [\, -2.41 - 0.50\,X_1 + 0.04\,X_2 + 0.51\,X_3 + 0.16\,X_4]\ X_5 \\ & + [\, 0.98 - 0.01\,X_1 + 0.11\,X_2 + 0.24\,X_4]\ X_5^2\,. \end{aligned} \qquad (13)$$

- in tetrahydrofuran-water (see appendix n°5):

$$\begin{aligned} \ln k'_T = \; & 0.28 + 0.33\,X_1 - 0.17\,X_2 - 0.64\,X_3 - 0.10\,X_4 \\ & + [\,-2.95 - 0.73\,X_1 + 0.12\,X_2 + 0.72\,X_3 + 0.04\,X_4]\;\; X_5 \\ & + [\,1.18 + 0.44\,X_1 - 0.09\,X_2 - 0.09\,X_4]\;\; X_5^2\,. \end{aligned} \qquad (14)$$

To check the weight of the different coefficients we performed a Pareto analysis. Plots of the contribution of every term are displayed in Figure 3.

The percentage effect P_i of every term i, is calculated through$^{(35)}$:

$$P_i = 100\left(\frac{b_i^2}{\sum b_j^2}\right) \qquad (15)$$

We can conclude from these plots:

- that the effect of the percentage of the modifier increases from methanol-water (66%) to acetonitrile-water (74%) and from acetonitrile-water (74%) to tetrahydrofuran-water (84%) according to the eluting strength.

- the terms X_5, X_1, X_{15} explain 91 % (65.65+ 15.26 + 10.01) of the lnk' variation with methanol-water as mobile phase,

- the terms X_5, X_1, X_3, X_{55} explain 92 % of the lnk' variation with acetonitrile-water as mobile phase,

- the terms X_5, X_{55}, X_3 explain 92 % of the lnk' variation with tetrahydrofuran-water as mobile phase.

The cavity term $V_i/100$ (X_1) is less important with tetrahydrofuran as modifier than it is with methanol and acetonitrile.

EXAMPLE OF THE MODEL APPLICATION

Let take as an example the separation of ethylbenzene from benzene with a column Hibar 100 RP-18 (Merck) endcapped (packed with Lichrospher 5µm) (250*4mm). Although this separation is not very exciting data are readily available and methodology may successfully be extended to any other pair of solutes.

The resolution $R_s$$^{(36)}$ can be written as :

$$R_s = \frac{\alpha-1}{\alpha+1}\cdot\frac{\bar{k}}{1+\bar{k}}\cdot\frac{\sqrt{N}}{2}\,. \qquad (16)$$

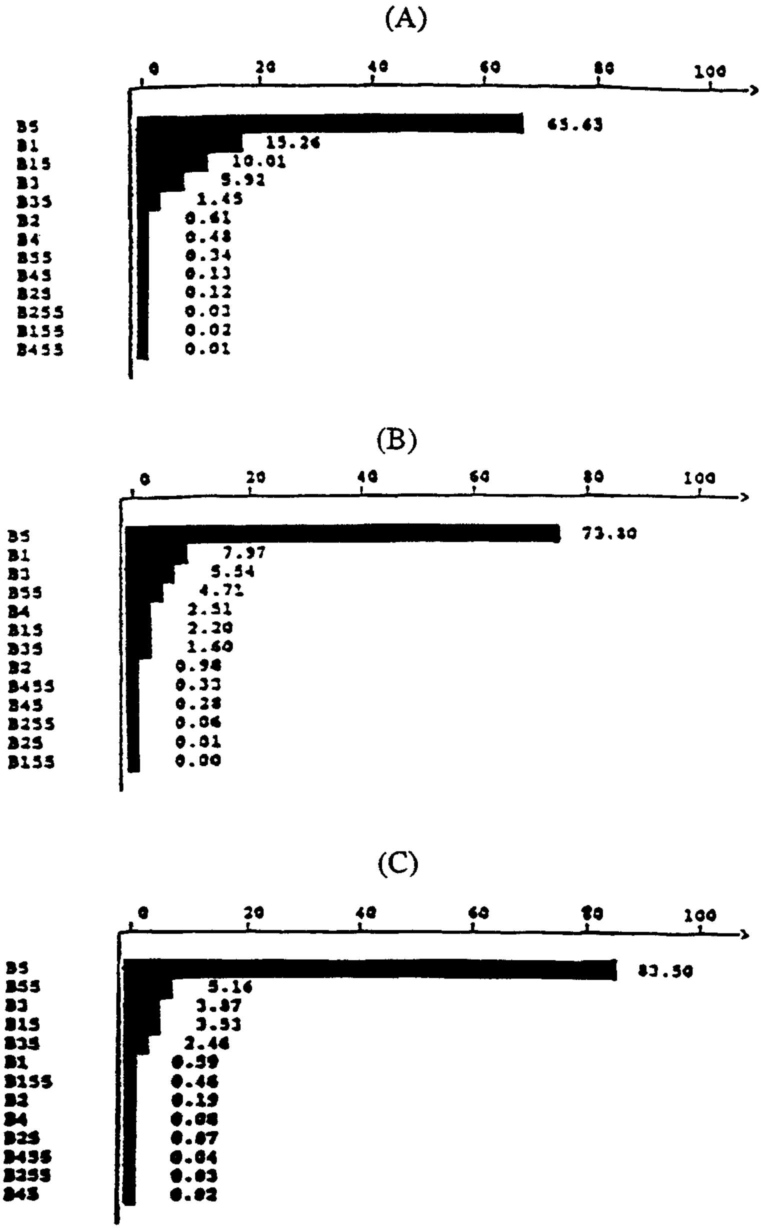


FIGURE 3. Pareto analysis

(A) methanol /water as eluent ; (B) acetonitrile /water as eluent ;

(C) tetrahydrofuran /water as eluent.

where :

α : the selectivity $\alpha_{12} = k_2/k_1$,

$\bar{k}$: the average capacity factor $\bar{k} = \frac{(k_1 + k_2)}{2}$,

N : the plate number.

If we decide to have an α value $\geq$ 2,2 with an optimal analysis time then we have to separate the two components in a range of $1 < \bar{k} < 5$ (37). The high value of α is selected according to the possibility of computer assisted sample pretreatment.

TABLE 6
Values of Natural Variables

		$V_i/100$	π^*	β	α
1	Benzene	0.491	0.59	0.10	0
2	Ethylbenzene	0.668	0.53	0.12	0

Table 6 can be converted as follows :

TABLE 7
Values of Coded Variables

		X_1	X_2	X_3	X_4
1	Benzene	-1	-0.75	-0.87	-1
2	Ethylbenzene	-0.32	-1	-0.79	-1

<u>In the case of methanol-water</u> :

Using equation (11) and values of Table (7) the lnk' of benzene in methanol-water binary mixture ($\ln k_{1M}$) becomes :

$$\ln k_{1M} = 2.582 + (2.255*(-1)) - (0.478*(-0.75)) - (0.994*(-0.87)) - (0.342*(-1)) + [-3.538 - (1.651*(-1)) + (0.359*(-0.75)) + (0.591*(-0.87)) + (0.305*(-1))] * X_5 + [0.453 - (0.125*(-1)) + (0.067*(-0.75)) - (0.049*(-1))] * X_5*X_5 . \quad (17)$$

then $\ln k_{1M} = 1.896 - 2.977\ X_5 + 0.576\ X_5^2$ (18)

and lnk' of ethylbenzene ($\ln k_{2M}$) becomes :

$$\ln k_{2M} = 3.456 - 4.133\ X_5 + 0.475\ X_5^2 . \quad (19)$$

As we look for value of $\alpha \geq 2.2$ then the following inequality must be fulfilled:

$$1.56 - 1.156\ X_5 - 0.101\ X_5^2 \geq \ln 2.2 \quad (20)$$

this inequality comes from :

$$\ln\alpha = \ln(k_{2M}/k_{1M}) = \ln k_{2M} - \ln k_{1M} = [\ 3.456 - 4.133\ X_5 + 0.475\ X_5^2] - [1.896 - 2.977\ X_5 + 0.576\ X_5^2] = 1.56 - 1.156\ X_5 - 0.101\ X_5^2 . \quad (21)$$

To respect this inequality X_5 must be :

$X_5 \leq 0.633$

or $X_5 \geq -12.055$

As $-1 \leq X_5 \leq +1$, one single solution is valid, which means that we have to choose the percent of methanol lower or equal to $(0.633*45+55) \simeq 83$ to get $\alpha \geq 2.2$.

At $X_5 = 0.633$, $\ln k_{1M} = 0.243$ and $\ln k_{2M} = 1.032$

then $\bar{k}_M = (k1M + k2M)/2 = 2.04$. (22)

<u>with acetonitrile-water</u> :

- lnk' of benzene according to the equation (13) and Table (7) becomes :

$$\ln k_{1A} = 1.282 - 2.546\ X_5 + 0.660\ X_5^2 . \quad (23)$$

- lnk' of ethylbenzene according to the equation (13) and Table(7) becomes:

$$\ln k_{2A} = 2.018 - 2.850\, X_5 + 0.628\, X_5^2. \qquad (24)$$

$$\Delta \ln k' = \ln k_{1A} - \ln k_{2A} = 0.736 - 0.304\, X_5 - 0.032\, X_5^2 \geq \ln 2.2.$$

then $X_5 \leq -0.175$.

At $X_5 = -0.175$, $\overline{k}_A = 9.19$.

<u>with tetrahydrofuran-water</u>:

- lnk' of benzene according to the equation (14) and Table(7) becomes :

$$\ln k_{1T} = 0.741 - 2.988\, X_5 + 0.888\, X_5^2. \qquad (25)$$

- lnk' of ethylbenzene according to the equation (14) and Table(7) becomes:

$$\ln k_{2T} = 0.951 - 3.448\, X_5 + 1.211\, X_5^2. \qquad (26)$$

$$\Delta \ln k' = \ln k_{1T} - \ln k_{2T} = 0.211 - 0.460\, X_5 + 0.322\, X_5^2 \geq \ln 2.2. \qquad (27)$$

then $X_5 \leq -0.804$.

At $X_5 = -0.804$, $\overline{k}_T = 65.69$.

TABLE 8.
Recapitulatory Table

$\alpha = 2.2$	Methanol/water (83/17) v/v	Acetonitrile/water (47/53) v/v	Tetrahydrofuran/water (19/81)v/v
$\overline{k}$	2.04	9.19	65.69

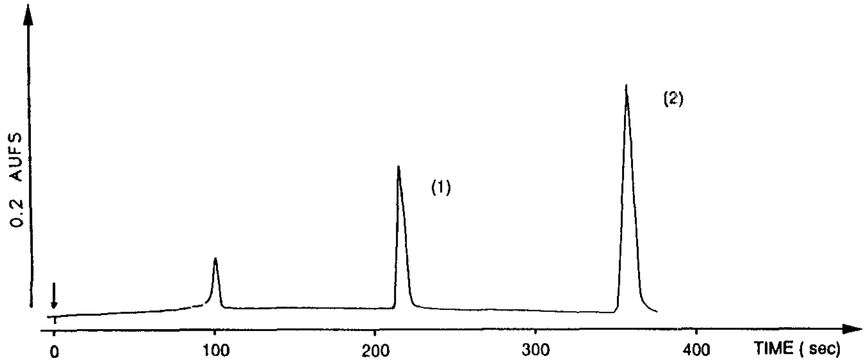


FIGURE 4. Separation of benzene from ethylbenzene

The separation of benzene from ethylbenzene was performed on :

- column : Hibar 100 RP-18 (Merck) endcapped (packed with Lichrospher 5μm) (250*4 mm)
- mobile phase : methanol/water (83/17) vol/vol
- injection volume : 20 μl
- flow-rate : 1 ml / min
- detection : UV detector - wavelength 254 nm.

(1) Benzene ; (2) Ethylbenzene.

Therefore, in order to obtain $\alpha \geq 2.2$ in minimal time, we have to select working with a 83/17 vol/vol methanol-water mixture. Chromatogram displayed in Fig 4 illustrates the validity of the approach since $\alpha_{exp} = 2.18$ and $\overline{k}_{exp} = 1.82$. These values are very close to those of Table 8.

EXPERIMENTAL

The liquid chromatography instrument was from Hitachi a 6000 A pump equipped with a Rheodyne 7125 sample loop (20 μl) and a fixed-wavelength detector (254 nm). Solvents were of Lichrosolv quality from Merck (Darmstadt, FRG). Water was distilled over potassium permanganate and purified by percolation through a RP 18 Lobar type column (Darmstadt, FRG). Solutes were from Sigma and used without further purification, dissolved in the selected mobile phase.

REFERENCES

1- R. F. Rekker, The Hydrophobic Fragmental Constant, Elsevier Publishing Company, Amsterdam, 1977.

2- C. Hansch, A. Leo, Substituent Constants for Correlation Analysis in Chemistry and Biology, John Wiley & Sons., Inc., New York, 1979.

3- A. Noble, J. Chromatogr., 642: 3 (1993).

4- R. M. Smith, C. M. Burr, J. Chromatogr., 550: 335 (1991).

5- R. M. Smith, C. M. Burr, J. Chromatogr., 485: 325 (1989).

6- R. M. Smith, C. M. Burr, J. Chromatogr., 481: 85 (1989).

7- R. M. Smith, C. M. Burr, J. Chromatogr., 481: 71 (1989).

8- R. M. Smith, C. M. Burr, J. Chromatogr., 475: 75 (1989).

9- R. M. Smith, C. M. Burr, J. Chromatogr., 475: 57 (1989).

10- A. Bondi, J. Phys. Chem., 68: 441 (1964).

11- L. B. Kier, L. H. Hall, Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York, 1976.

12- J. F. Schabron, R. J. Hurtubise, H. F. Silver, Anal. Chem., 49: 2253 (1977).

13- R. J. Hurtubise, T. W. Allen, H. F. Silver, J. Chromatogr., 235: 517 (1982).

14- R. W Taft, J. L. M Abboud, M. J. Kamlet, M. H. Abraham, J. Solution Chem., 14: 153 (1985).

15- M. J. Kamlet, R. M. Doherty, J. L. M. Abboud, R. W. Taft, M. H. Abraham, Chemtech., 16: 566 (1986).

16- M. J. Kamlet, R. M. Doherty, M. H. Abraham, Y. Marcus, R. W. Taft, J. Phys. Chem., 92: 5244 (1988).

17- J. C. McGowan, J. Appl. Chem. Biotechnol., 34A: 38 (1984).

18- P. C. Sadek, P. W. Carr, R. M. Doherty, M. J. Kamlet, R. W. Taft, M. H. Abraham, Anal. Chem., 57: 2971 (1985).

19- M. J. Kamlet, M. H. Abraham, P. W. Carr, R. M. Doherty, R. W. Taft, J. Chem. Soc. Perkin Trans. II., 2087 (1988).

20- R. Phan-Tan-Luu, D. Bazard, R. Flayeux, J. M. Lehuede, M. Parisot, T. Vinh, M. Moll, Industries Alimentaires et Agricoles, 441 (1979).

21- R. M. Claramunt, R. Gallo, J. Elguero, D. Mathieu, R. Phan-Tan-Luu, J. Chim. Phys., 78: 805 (1981).

22- E. Fargin, M. Sergent, D. Mathieu, R. Phan-Tan-Luu, Bio-sciences., 4 (1985).

23- K. Jinno, K. Kawasaki, J. Chromatogr., 316: 1 (1984).

24- R. F. Gunst, R. L. Mason, Regression Analysis and its Application, Marcel Dekker, Inc. New York and Basel, 1980.

25- P. J. Schoenmakers, H. A. H. Billiet, R. Tijssen, L. De Galan, J. Chromatogr., 149: 519 (1978).

26- P. J. Schoenmakers, H. A. H. Billiet, L. De Galan, J. Chromatogr., 185: 179 (1979).

27- K. A. Dill, J. Phys. Chem., 91: 1980 (1987).

28- P. Jandera, J. Chromatogr., 314: 13 (1984).

29- M. C. Pietrogrande, C. Bighi, P. A. Borea, F. Dondi, J. Chromatogr., 522: 37 (1990).

30- Hanfa Zou, Yukui Zhang, Peichang Lu, J. Chromatogr., 522: 49 (1990).

31- S. Chatterjee, B. Price, Regression Analysis by Example, John Wiley & Sons., Inc., New York, 1977.

32- J. G. Hahn, W. Q. Meeker, P. I. Feder, J. Qual. Techn., 8: 140 (1976).

33- D. W. Marquardt, Technometrics, 12: 591 (1970).

34- R. D. Snee, J. Qual. Techn., 5: 67 (1973).

35- P. D. Haaland, Experimental Design in Biotechnology, Marcel Dekker, Inc. New York and Basel, 1989.

36- P. J. Schoenmakers, Optimization of Chromatographic Selectivity, Journal of Chromatography Library - Volume 35, Elsevier, 1987.

37- L. R. Snyder, J. J. Kirkland, Introduction to Modern Liquid Chromatography, John Wiley & Sons., Inc., New York, 1979.

APPENDIX n° 1

CHARACTERISTICS OF THE CHOSEN MATRIX

CHARACTERISTICS

Log Determinant (X'X) = 15.813692
Log Determinant (M) = -12.312227
Log Determinant (M)** 1/NbrCoeff = -0.820815
Maximal variance fonction = 0.657
Trace (X'X)-1 = 3.514
G (%) efficiency = 30.451

Name	Variance coefficient	Inflation factor	Name	Variance coefficient	Inflation factor
b0	0.1124	0.000			
b1	0.1607	2.746	b2	0.2124	6.819
b3	0.2275	8.531	b4	0.1283	3.826
b5	0.0926	3.471	b51	0.1324	1.852
b52	0.1749	2.910	b53	0.1874	3.568
b54	0.1056	2.870	b55	0.2645	3.471
b551	0.3782	3.467	b552	0.4998	6.922
b553	0.5353	8.587	b554	0.3018	5.120

APPENDIX n°2

CHARACTERISTICS OF THE MATRIX AFTER REDUCING THE MODEL

CHARACTERISTICS

Log Determinant (X'X) = 15.542320
Log Determinant (M) = -10.708538
Log Determinant (M)** 1/NbrCoeff = -0.764896
Maximal variance fonction = 0.642
Trace (X'X)-1 = 2.338
G (%) efficiency = 29.065

Name	Variance coefficient	Inflation factor	Name	Variance coefficient	Inflation factor
b0	0.1057	0.000			
b1	0.1606	2.744	b2	0.1412	4.532
b3	0.0937	3.513	b4	0.1051	3.135
b5	0.0926	3.471	b51	0.1324	1.852
b52	0.1749	2.910	b53	0.1874	3.568
b54	0.1056	2.870	b55	0.2375	3.117
b551	0.3778	3.463	b552	0.2149	2.976
b554	0.2091	3.548			

(continued)

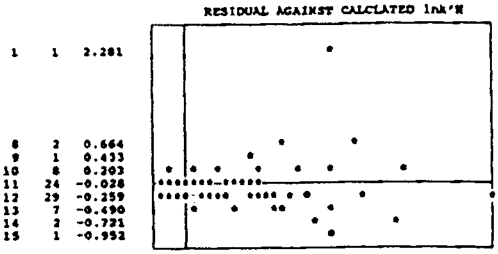
APPENDIX n°2 (continued)

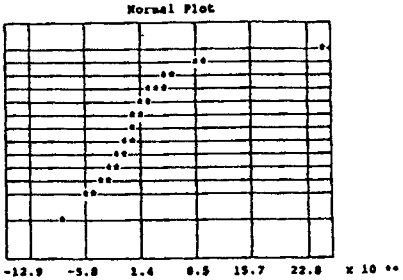
LNK' VERSUS THE FIVE FACTORS : RELATIONSHIP WITH METHANOL-WATER

Source	Sums of squares	Degrees of freedom	Mean sum of squares	Ratio	SIGNIF
REGRESSION	510.08116	13	39.23701	227.72	***
RESIDUAL	10.51065	61	0.17231		
TOTAL	520.59181	74			

Estimation of standard deviation	:	0.4151
Multiple Regression Coeff.(R2)	:	0.9798
Multiple Regression Coeff.(R2A)	:	0.9755
Degrees of Freedom	:	61

VAR	COEFFICIENT	INFLATION FACTOR	STANDARD DEVIATION	t Exp.	Signif
b0	2.58154		0.13492	19.133	***
b1	2.25531	2.7443	0.16637	13.556	***
b2	-0.47752	4.5324	0.15598	-3.061	**
b3	-0.99405	3.5129	0.12705	-7.824	***
b4	-0.34201	3.1354	0.13456	-2.542	**
b5	-3.53797	3.4713	0.12629	-28.014	***
b51	-1.65112	1.8520	0.15103	-10.933	***
b52	0.35862	2.9104	0.17362	2.066	*
b53	0.59075	3.5685	0.17968	3.288	***
b54	0.30468	2.8696	0.13490	2.259	*
b55	0.45298	3.1169	0.20228	2.239	*
b551	-0.12505	3.4634	0.25514	-0.490	31.6%
b552	0.06684	2.9757	0.19242	0.347	36.5%
b554	-0.04856	3.5485	0.18982	-0.256	39.7%





$$\ln k'_M = 2.582 + 2.255\,X_1 - 0.478\,X_2 - 0.994\,X_3 - 0.342\,X_4 + [-3.538 - 1.651\,X_1 + 0.359\,X_2 + 0.591\,X_3 + 0.305\,X_4]\,X_5 + [0.453 - 0.125\,X_1 + 0.067\,X_2 - 0.049\,X_4]\,X_5^2 \quad (11)$$

APPENDIX n° 3

Source	Sums of squares	Degrees of freedom	Mean sum of squares	Ratio	SIGNIF
REGRESSION	480.40736	13	36.95441	484.96	***
RESIDUAL	4.57207	60	0.07620		
TOTAL	484.97943	73			

Estimation of standard deviation: 0.2760
Multiple Regression Coeff.(R2) : 0.9906
Multiple Regression Coeff.(R2A): 0.9885
Degrees of Freedom .... : 60

VAR	COEFFICIENT	INFLATION FACTOR	STANDARD DEVIATION	t Exp.	Signif
b0	2.60701		0.08977	29.040	***
b1	2.25115	2.7439	0.11064	20.347	***
b2	-0.42275	4.5364	0.10391	-4.068	***
b3	-1.07753	3.5567	0.08502	-12.674	***
b4	-0.32173	3.1180	0.08952	-3.594	***
b5	-3.54414	3.3777	0.08399	-42.197	***
b51	-1.65467	1.8266	0.10044	-16.475	***
b52	0.20811	2.9688	0.11671	1.783	*
b53	0.75772	3.6575	0.12098	6.263	***
b54	0.16782	2.8451	0.09104	1.843	*
b55	0.40823	3.0610	0.13462	3.033	**
b551	-0.11319	3.4407	0.16967	-0.667	25.7%
b552	0.10783	2.9681	0.12805	0.842	20.4%
b554	0.04772	3.4660	0.12671	0.377	35.4%

APPENDIX n° 4

LNK' VERSUS THE FIVE FACTORS : RELATIONSHIP WITH ACETONITRILE-WATER:

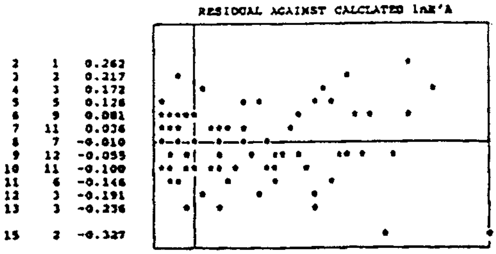
Source	Sums of squares	Degrees of freedom	Mean sum of squares	Ratio	SIGNIF
REGRESSION	262.53885	13	20.19530	1051.85	***
RESIDUAL	1.17119	61	0.01920		
TOTAL	263.71004	74			

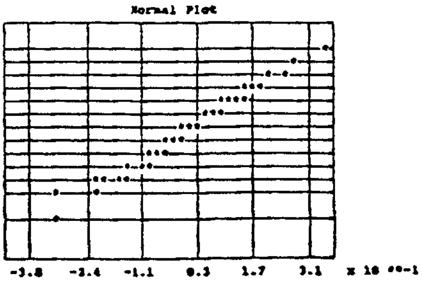
Estimation of standard deviation: 0.1386
Multiple Regression Coeff.(R2) : 0.9956
Multiple Regression Coeff.(R2A): 0.9946
Degrees of Freedom .... : 61

(continued)

APPENDIX n° 4 (continued)

VAR	COEFFICIENT	INFLATION FACTOR	STANDARD DEVIATION	t Exp.	Signif
b0	1.01319		0.04504	22.496	***
b1	1.04463	2.7443	0.05554	18.810	***
b2	-0.34323	4.5324	0.05207	-6.592	***
b3	-0.66513	3.5129	0.04241	-15.683	***
b4	-0.47404	3.1354	0.04492	-10.553	***
b5	-2.41250	3.4713	0.04216	-57.225	***
b51	-0.49802	1.8520	0.05041	-9.879	***
b52	0.04003	2.9104	0.05796	0.691	25.0%
b53	0.50568	3.5685	0.05998	8.431	***
b54	0.15913	2.8696	0.04503	3.534	***
b55	0.97626	3.1169	0.06752	14.458	***
b551	-0.00699	3.4634	0.08517	-0.082	46.6%
b552	0.10783	2.9757	0.06423	1.679	*
b554	0.24260	3.5485	0.06337	3.829	***





$$
\begin{aligned}
\ln k'_A = \; & 1.01 + 1.05\,X_1 - 0.34\,X_2 - 0.67\,X_3 - 0.47\,X_4 \\
& + [\,-2.41 - 0.50\,X_1 + 0.04\,X_2 + 0.51\,X_3 + 0.16\,X_4]\;\; X_5 \\
& + [\,0.98 - 0.01\,X_1 + 0.11\,X_2 + 0.24\,X_4]\;\; X_5^2\,. \qquad (13)
\end{aligned}
$$

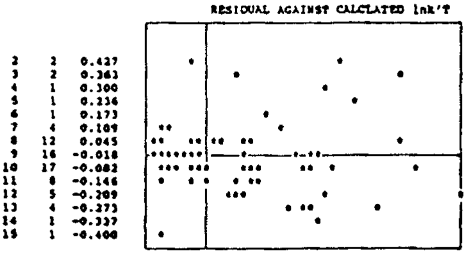
APPENDIX n° 5
LNK' VERSUS THE FIVE FACTORS : RELATIONSHIP WITH THF-WATER :

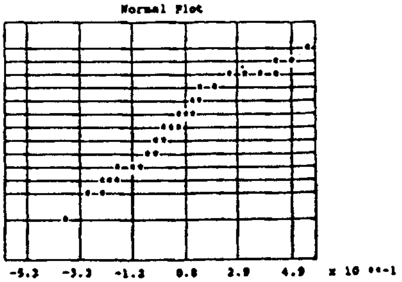
Source	Sums of squares	Degrees of freedom	Mean sum of squares	Ratio	SIGNIF
REGRESSION	316.32670	13	24.33282	777.37	***
RESIDUAL	1.90938	61	0.03130		
TOTAL	318.23607	74			

Estimation of standard deviation: 0.1769
Multiple Regression Coeff.(R2) : 0.9940
Multiple Regression Coeff.(R2A): 0.9927
Degrees of Freedom .... : 61

APPENDIX n°5 (continued)

VAR	COEFFICIENT	INFLATION FACTOR	STANDARD DEVIATION	t Exp.	Signif
b0	0.28013		0.05751	4.871	***
b1	0.32648	2.7443	0.07091	4.604	***
b2	-0.17399	4.5324	0.06648	-2.617	**
b3	-0.63949	3.5129	0.05415	-11.809	***
b4	-0.09687	3.1354	0.05735	-1.689	*
b5	-2.95239	3.4713	0.05383	-54.848	***
b51	-0.72632	1.8520	0.06437	-11.283	***
b52	0.11690	2.9104	0.07400	1.580	5.8%
b53	0.72102	3.5685	0.07658	9.415	***
b54	0.04318	2.8696	0.05750	0.751	23.1%
b55	1.17525	3.1169	0.08622	13.631	***
b551	0.44470	3.4634	0.10874	4.089	***
b552	-0.08820	2.9757	0.08201	-1.075	14.3%
b554	-0.09157	3.5485	0.08091	-1.132	13.1%





$$
\begin{aligned}
\ln k'_T = {} & 0.28 + 0.33\, X_1 - 0.17\, X_2 - 0.64\, X_3 - 0.10\, X_4 \\
& + [\,-2.95 - 0.73\, X_1 + 0.12\, X_2 + 0.72\, X_3 + 0.04\, X_4]\ X_5 \\
& + [\,1.18 + 0.44\, X_1 - 0.09\, X_2 - 0.09\, X_4]\ X_5^2\,. \qquad (14)
\end{aligned}
$$

Received: October 10, 1993
Accepted: February 10, 1994