Journal of Chromatography, 550 (1991) 335–356
Elsevier Science Publishers B.V., Amsterdam

CHROMSYMP. 2180

Retention prediction of analytes in reversed-phase high-performance liquid chromatography based on molecular structure

VI.[a] Disubstituted aromatic compounds

ROGER M. SMITH* and CHRISTINA M. BURR
Department of Chemistry, Loughborough University of Technology, Loughborough, Leics. LE11 3TU (UK)

ABSTRACT

As part of the development of a model for retention prediction based on molecular structure, the effects of interactions between substituents on the retention of aromatic analytes in reversed-phase high-performance liquid chromatography were examined, using 73 *ortho*-, *meta*- and *para*-disubstituted aromatic compounds. The interactions can be expressed empirically as eluent-dependent interaction terms. A more general expression was also examined that includes elements for electronic interactions, dependent on the Hammett constants of the substituents and their susceptibility, *ortho*-steric effects and hydrogen bonding.

INTRODUCTION

A number of computer-based methods have been devised to aid the development of separations in high-performance liquid chromatography (HPLC) [1]. The most common approach has been optimization methods, which rely on the combination of experimental observation with calculations, and require no knowledge of the structure of the analyte. In contrast, relatively few methods have been proposed that estimate the retention from the molecular structure of the analyte, although a number of groups have worked within restricted groups of compounds [2,3], or on more specialized separations, such as ion-exchange chromatography [4,5]. Close correlations have also been drawn between structure-based molecular connectivity calculations and retentions [6]. The structure or partial structure of the analyte is frequently known or the structural differences between related compounds, such as isomers or metabolites, are known and it should be possible to use this information as the basis of a prediction system.

[a] For Part V, see ref. 10.

0021-9673/91/$03.50

The aim of this project has been to develop a retention prediction method for reversed-phase liquid chromatography based on the molecular structure of the analyte [7]. The intention is to calculate the retention index as the summation of a series of terms:

$$I = I_P + I_{S,R} + \sum I_{S,Al-X} + \sum I_{S,Ar-X} + \sum I_{I,Y-Z} \quad (1)$$

These represent the retention index of a parent compound (I_P), a contribution for saturated alkyl chains ($I_{S,R}$), contributions for substituents on saturated aliphatic carbons ($I_{S,Al-X}$), contributions for aromatic substituents ($I_{S,Ar-X}$) and terms to account for any interactions between substituents ($I_{I,Y-Z}$) caused by electronic, hydrogen bonding and steric effects. Each of these terms will be sensitive to eluent composition and the organic modifier in the eluent and will be related to the percentage of modifier (x) using a quadratic equation ($I = ax^2 + bx + c$). Benzene was selected as the parent compound as its substituted derivatives can be readily detected. Retention indices based on the alkyl aryl ketones were used as the basis of the study as they are more robust than capacity factors and can be more readily transferred between systems. So far the terms for individual aromatic [7] and aliphatic [8] substituents and structural isomerization [9] have been determined. The coefficients of the regression equations are held in a database, which can be interrogated by an expert system program (CRIPES; Chromatographic Retention Index Prediction Expert System) for the calculation of predicted retention indices [10].

This model for retention prediction resembles the methods used by Hansch and others to calculate octanol–water partition coefficients (log P) [11] and in the early stages of this study it was shown that the retention substituent indices for groups on an aromatic ring were closely related to the Hansch substituent contributions (π) [7]. Hansch found that for disubstituted aromatic compounds containing polar groups, simple summation of the π terms to calculate log P was not very successful [12]. This was assumed to be due to interactions between the substituents and led initially to individual sets of π_Y values for each parent substituent. However, this approach was clearly unsatisfactory as it leads to multiple sets of π values and a proliferations of data.

For the *meta* and *para* isomers it was found that the difference between the π value of a substituent X with benzene as the parent compound and that with phenol as the parent could be described using the Hammett constant σ of the substituent [12]. This led to a more general equation, which attempted to quantify the effect of a group Y on the Hansch constant of a substituent X in terms of their "susceptibility" constants (ρ) and Hammett constants (σ) [12–14].

$$\pi_{X(PhY)} - \pi_{X(PhH)} = \rho_Y\sigma_X + \rho_X\sigma_Y \quad (2)$$

where ρ_X and ρ_Y are the susceptibilities of X and Y to the modifying effects of Y and X, respectively. Values for the susceptibility constants were derived experimentally using multiple regression analysis. This approach was extended to *ortho* substituents by adding additional terms to account for the proximity effects [14]:

$$\Sigma_{o\text{-}XPhY} = a\pi_{XPhY} + \rho_Y\sigma^0_{o-X} + \rho_X\sigma^0_{o\text{-}Y} + f_YF_X + f_XF_Y + \delta_YE^X_S + \delta_XE^Y_S + c \quad (3)$$

where $a\pi$ = the additive π, E_S = Taft steric effect value, F = Swain–Lupton field-effect constant, f = the susceptibility for the F constants and δ the susceptibility for the Taft E_S values. The E_S term was found to be insignificant for phenols and anilines (in the absence of intramolecular hydrogen bonding).

In a similar study, Leo [15,16] developed a simplified model to facilitate the rapid estimation of the interaction terms. The sigma/rho interaction term (F_σ) was calculated in the same way as by Fujita [14] but closely related substituents were assigned a common σ value and the same σ and ρ values were used for *ortho, meta* and *para* substituents. He also considered that for many groups one factor or the other was so small as to be safely ignored. The substituents were divided into three classes, either inducers (electron withdrawing) with $\rho = 0$ (*e.g.*, CN, NO_2, halogen), responders (electron releasing) with $\sigma = 0$ (*e.g.*, OH, NH_2) or bi-directional with $\rho > 0$ and $\sigma > 0$ (*e.g.*, CHO, CO_2CH_3 and OR), whose overall effect would be governed by the second substituent present in the compound. Additional terms were included to account for intramolecular hydrogen bonding (F_{HB}), the negative *ortho* effect (F_o) and the presence of alkyl–aryl systems ($F_{\alpha\varphi}$). Linear regression analysis gave the following correlation with experimental partition coefficients:

$$\log P = \Sigma\pi + F_\sigma - 0.29F_o + 0.63F_{HB} - 0.15F_{\alpha\varphi} \quad (4)$$

in which

$$F_\sigma = \rho_Y\sigma_X + \rho_X\sigma_Y$$

With the exception of the electronic effects each of the terms was quantized taking values 0, 1, 2, etc. Although a single intramolecular hydrogen bonding term was suggested, this was found to be insufficient to account for the observed hydrogen bonding effects between *ortho*-hydroxyl and amide groups.

This paper describes the examination of the retentions of a number of isomeric substituted toluenes and phenols in methanol–buffer and acetonitrile–buffer eluents. These studies have led to a set of coefficients for empirically based interaction indices (I_I), which have been used in the retention prediction system. However, because the applicability of these terms is limited to these specific substituent pairs, the results were also examined to determine if a model similar to that developed by Leo (eqn. 4) would be appropriate, which would have general applicability to any pair of substituents.

EXPERIMENTAL

Chemicals, equipment and procedures were as described previously [7].

Calculation of interaction indices

The retention index increments for each interaction (δI) were calculated as the difference between the retention index of the disubstituted compound (I_{Expt}) and the calculated retention index (I_{Sum}) for the same eluent based on the summation of the parent index value of benzene and the individual contributions for the substituents derived earlier [7].

DISCUSSION

In order to study the interactions, 73 *ortho*-, *meta*- and *para*-substituted toluenes and phenols were examined. The former compounds would be expected to show only minor effects, mainly due to steric or electronic interactions, whereas the phenols are likely to demonstrate stronger electronic effects and many are capable of intramolecular hydrogen bonding. The capacity factors of the compounds were determined in a range of eluents from 40–80% methanol–pH 7 buffer and 30–80% acetonitrile–pH 7 buffer (Tables I and II). In each instance the retentions of the homologous alkyl aryl

TABLE I

CAPACITY FACTORS OF SUBSTITUTED TOLUENES

Compound	Capacity factor (k')										
	Methanol (%)					Acetonitrile (%)					
	40	50	60	70	80	30	40	50	60	70	80
2-Bromotoluene	105.99	39.60	15.27	6.54	2.63	72.19	24.95	10.61	5.24	2.88	1.67
3-Bromotoluene	106.98	39.79	15.39	6.48	2.59	75.32	25.59	10.59	5.20	2.83	1.62
4-Bromotoluene	104.87	38.94	15.02	6.34	2.54	74.98	25.27	10.53	5.16	2.80	1.61
2-Chlorotoluene	83.81	32.24	12.82	5.48	2.31	64.45	21.19	9.16	4.73	2.56	1.50
3-Chlorotoluene	–	31.77	12.72	5.42	2.27	65.48	21.29	9.13	4.68	2.51	1.45
4-Chlorotoluene	76.31	32.55	12.47	5.26	2.20	65.56	21.45	9.07	4.64	2.48	1.44
1,2-Dimethylbenzene	60.72	26.61	11.34	5.11	2.25	54.74	18.27	8.16	4.16	2.34	1.37
1,3-Dimethylbenzene	70.30	29.85	12.57	5.53	2.38	63.40	20.43	8.99	4.54	2.52	1.47
1,4-Dimethylbenzene	72.19	30.35	13.17	5.63	2.41	63.64	20.73	9.11	4.60	2.56	1.47
2-Methylacetophenone	13.35	5.76	2.92	1.55	0.87	12.59	6.02	2.96	1.77	1.07	0.76
3-Methylacetophenone	15.04	5.81	2.98	1.56	0.84	12.50	5.91	2.89	1.73	1.05	0.75
4-Methylacetophenone	14.81	5.82	2.94	1.54	0.83	11.70	5.56	3.04	1.66	1.02	0.74
2-Methylanisole	34.63	14.72	7.06	3.43	1.68	33.85	14.67	7.01	3.32	1.82	1.18
3-Methylanisole	28.90	11.76	5.80	2.92	1.47	26.81	11.94	5.84	2.84	1.58	1.04
Methyl 2-methylbenzoate	28.70	11.62	5.02	2.47	1.22	25.82	8.29	4.10	2.29	1.37	0.88
Methyl 3-methylbenzoate	32.83	12.72	5.37	2.58	1.24	27.57	8.60	4.19	2.35	1.40	0.89
Methyl 4-methylbenzoate	32.90	12.58	5.35	2.58	1.24	26.78	8.44	4.16	2.32	1.40	0.90
2-Nitrotoluene	16.72	7.60	3.53	1.83	0.94	21.20	7.15	3.49	1.94	1.14	0.71
3-Nitrotoluene	19.99	8.93	4.24	2.14	1.08	24.82	8.41	3.92	2.14	1.25	0.77
4-Nitrotoluene	18.53	8.45	3.96	2.02	1.05	21.72	7.67	3.74	2.06	1.21	0.75
2-Phenyltoluene	–	105.96	34.17	14.71	4.21	–	77.11	22.18	9.09	4.12	2.31
3-Phenyltoluene	–	125.14	38.49	13.72	4.87	–	83.16	23.70	9.60	4.32	2.40
4-Phenyltoluene	–	141.24	42.17	11.65	5.37	–	87.94	24.83	10.02	4.49	2.49
2-Tolualdehyde	10.70	4.87	2.47	1.41	0.84	10.73	5.43	2.89	1.70	1.11	0.58
3-Tolualdehyde	11.00	4.88	2.47	1.37	0.83	11.26	5.62	2.98	1.73	1.14	0.58
4-Tolualdehyde	10.16	4.57	2.33	1.32	0.81	10.18	5.15	2.77	1.65	1.10	0.57
2-Toluamide	1.43	0.97	0.49	0.39	0.28	1.35	0.74	0.52	0.41	0.33	0.34
3-Toluamide	2.40	1.19	0.67	0.48	0.33	1.70	0.92	0.62	0.46	0.36	0.33
4-Toluamide	2.44	1.14	0.66	0.47	0.33	1.90	1.01	0.61	0.47	0.36	0.35
2-Toluidine	3.39	1.91	1.09	0.72	0.45	4.65	2.30	1.44	0.97	0.65	0.48
3-Toluidine	3.52	1.98	1.10	0.71	0.44	4.87	2.33	1.43	0.96	0.64	0.47
4-Toluidine	3.77	2.04	1.12	0.70	0.45	4.85	2.30	1.43	0.96	0.66	0.48
2-Toluonitrile	10.81	5.28	2.32	1.29	0.71	11.43	5.08	2.67	1.58	0.98	0.65
3-Toluonitrile	11.88	4.82	2.52	1.36	0.73	13.03	5.65	2.91	1.68	1.04	0.66
4-Toluonitrile	11.39	5.04	2.43	1.31	0.70	12.21	5.33	2.77	1.61	0.99	0.65

ketones, acetophenone–heptanophenone, were also measured and these were used to calculate the retention indices (I) of the analytes (Tables III and IV). A few of the compounds were eluted too rapidly for accurate measurement ($k' < 0.2$) and the corresponding indices are given in parentheses as they may be unreliable [10]. In some instances, such as with the hydroxybenzamides and dihydroxybenzenes, the retention index scale required considerable extrapolation and these values may also be less reliable.

Although the capacity factors changed significantly with eluent composition, the retention indices were usually relatively constant across the composition range. For most of the substituents the retention indices of the *meta* and *para* isomers were similar (± 50 units) but frequently that of the *ortho* isomer was significantly different (up to 400 units). Both the capacity factors and retention indices for 4-nitrophenol (pK_a = 7.1) [17] appeared to be abnormally low, particularly in methanol–buffer eluents, and it was suspected that this acidic phenol was significantly ionized in the mobile phase. Subsequently it appeared that 4-hydroxybenzaldehyde (pK_a = 7.6), 2-cyanophenol (pK_a = 6.9), 4-cyanophenol (pK_a = 7.7–7.9) and 2-nitrophenol (pK_a = 7.2) [17] might also be partially ionized, particularly with high proportions of methanol in the eluent. The results from these compounds were therefore regarded as potential outliers and were excluded from any correlation studies.

Because of the close relationship between octanol–water partition coefficients ($\log P$) and $\log k'$, which has often been used in QSAR studies [18], the retention indices in methanol–buffer (60:40) were compared with reported $\log P$ values [11,15] (Tables III and IV, Fig. 1). Although there was a good correlation for the substituted toluenes, there were systematic differences between the toluenes and many of the phenols. This difference agrees with the contribution of the phenolic group to the retention index ($I_{S,ArOH}$), which was about 90 units more negative than that predicted [7] from a linear relationship between I_S and π. Similar marked differences for ionizable compounds were also reported by Miyake *et al.* [19].

However, phenols with an *ortho*-carbonyl substituent capable of strong intramolecular hydrogen bonding ($COCH_3$, CO_2CH_3, $CONH_2$ and NH_2) behaved similarly to the toluenes, suggesting that the effect of the hydroxy group was largely masked in these compounds. The retentions of the compounds suspected of being ionized were all lower than the correlation curve for the non-ionized phenols.

Interaction increments

In order to study the interactions between the substituents, estimated retention indices (I_{Sum}) for each eluent were calculated by the summation of the previously determined parent index values for benzene and the substituent indices (I_S) for the individual groups [7] (Table V). These sums do not contain a contribution for interactions between the substituents so that the *ortho*, *meta* and *para* isomers have the same values. The interaction increments for each pair of substituents were then calculated as $\delta I = I_{Expt} - I_{Sum}$ (Tables VI and VII).

In order to represent these values in the retention prediction program [10], the relationship between the increment for each pair of substituents and the eluent composition was expressed as a quadratic expression (Tables VIII and IX). These coefficients can then be used to calculate the interaction increments ($I_I = ax^2 + bx + c$). For many substituted toluenes the interaction increments for each eluent

TABLE II

CAPACITY FACTORS OF SUBSTITUTED PHENOLS[a]

Compound	Capacity factor (k')										
	Methanol (%)					Acetonitrile (%)					
	40	50	60	70	80	30	40	50	60	70	80
2-Aminophenol	0.78	0.66	0.42	0.30	0.26	1.08	0.79	0.63	0.54	0.42	0.44
3-Aminophenol	0.35	0.28	0.25	(0.19)	(0.18)	0.63	0.52	0.38	0.30	0.23	0.26
4-Aminophenol	0.25	0.22	(0.19)	(0.16)	0.21	0.40	0.38	0.29	0.24	0.20	0.26
2-Bromophenol	6.12	3.34	1.70	0.86	0.45	5.82	3.06	1.60	0.97	0.61	0.45
3-Bromophenol	9.38	4.82	2.17	1.11	0.56	7.85	3.77	1.84	1.08	0.66	0.48
4-Bromophenol	8.80	4.56	2.14	1.10	0.57	7.49	3.60	1.77	1.05	0.65	0.49
2-Chlorophenol	4.86	2.72	1.36	0.74	0.42	4.69	2.75	1.43	0.89	0.56	0.44
3-Chlorophenol	7.38	3.95	1.85	0.93	0.51	6.98	3.26	1.64	0.99	0.61	0.47
4-Chlorophenol	6.94	3.65	1.76	0.91	0.51	6.39	3.04	1.57	0.94	0.59	0.45
1,2-Dihydroxybenzene	0.96	0.64	0.43	0.31	0.25	0.85	0.78	0.53	0.42	0.31	0.41
1,3-Dihydroxybenzene	0.53	0.37	0.27	(0.19)	(0.18)	0.55	0.55	0.38	0.29	0.22	0.25
1,4-Dihydroxybenzene	0.31	0.27	0.20	(0.18)	(0.18)	0.35	0.43	0.32	0.25	0.21	0.24
2-Hydroxyacetophenone	9.02	4.33	2.21	1.25	0.84	7.57	4.51	2.16	1.52	0.84	0.65
3-Hydroxyacetophenone	2.02	1.02	0.58	0.39	0.34	1.70	1.31	0.67	0.55	0.29	0.32
4-Hydroxyacetophenone	1.55	0.78	0.44	0.29	(0.19)	1.15	0.88	0.53	0.46	0.27	0.29
2-Hydroxybenzaldehyde	4.99	2.52	1.41	0.83	0.53	6.86	5.56	2.39	1.34	0.94	0.56
3-Hydroxybenzaldehyde	1.64	0.89	0.54	0.37	0.25	1.77	1.11	0.74	0.50	0.35	(0.19)
4-Hydroxybenzaldehyde	1.00	0.51	0.27	(0.14)	(0.07)	1.20	0.78	0.54	0.40	0.29	(0.16)

2-Hydroxybenzamide	1.57	0.84	0.50	0.31	0.20	1.52	0.95	0.76	0.47	0.35	0.21
4-Hydroxybenzamide	0.32	0.22	(0.16)	(0.13)	(0.11)	0.27	0.23	(0.19)	(0.17)	(0.14)	(0.01)
2-Hydroxybenzonitrile	1.31	0.74	0.34	(0.18)	–	1.51	1.02	0.66	0.48	0.30	0.22
3-Hydroxybenzonitrile	2.17	1.26	0.68	0.42	0.21	2.50	1.47	0.87	0.58	0.38	0.33
4-Hydroxybenzonitrile	1.48	0.89	0.47	0.26	(0.07)	1.74	1.18	0.73	0.49	0.34	0.26
2-Methoxyphenol	2.95	1.49	0.92	0.59	0.37	3.12	1.90	1.12	0.77	0.51	0.41
3-Methoxyphenol	2.63	1.29	0.78	0.50	0.31	2.73	1.62	0.97	0.64	0.42	0.32
4-Methoxyphenol	1.93	0.99	0.63	0.42	0.29	2.05	1.30	0.82	0.56	0.38	0.29
Methyl 2-hydroxybenzoate	19.60	10.09	4.65	2.31	1.16	14.59	7.29	3.41	2.00	1.18	0.82
Methyl 3-hydroxybenzoate	3.72	1.96	0.99	0.58	0.34	3.10	1.67	0.95	0.61	0.41	0.36
Methyl 4-hydroxybenzoate	3.39	1.88	0.88	0.49	0.25	2.44	1.48	0.88	0.61	0.40	0.33
2-Methylphenol	4.75	2.76	1.42	0.78	0.47	5.26	2.55	1.41	0.89	0.58	0.41
3-Methylphenol	4.41	2.44	1.29	0.72	0.44	4.45	2.28	1.28	0.81	0.53	0.25
4-Methylphenol	4.86	2.51	1.31	0.73	0.45	5.10	2.28	1.28	0.82	0.54	0.24
2-Nitrophenol	3.56	1.96	0.91	0.40	(0.08)	5.27	2.59	1.57	1.07	0.65	0.36
3-Nitrophenol	3.38	1.96	1.00	0.53	0.23	4.23	2.03	1.12	0.71	0.45	0.35
4-Nitrophenol	1.27	0.69	0.29	(0.13)	–	2.12	1.06	0.67	0.49	0.28	(0.19)
2-Phenylphenol	33.87	10.73	4.48	1.97	0.90	31.39	10.51	4.15	1.98	1.04	0.67
3-Phenylphenol	35.87	11.35	4.53	1.92	0.92	27.91	9.01	3.63	1.76	0.93	0.62
4-Phenylphenol	35.74	11.52	4.12	1.98	0.96	27.59	8.86	3.36	1.72	0.99	0.63

[a] Values in parentheses in this and subsequent tables are for analytes with capacity factors less than 0.20 and may be unreliable.

ABLE III

ETENTION INDICES OF SUBSTITUTED TOLUENES

ompound	Retention index (I)											Log P^a
	Methanol (%)					Acetonitrile (%)						
	40	50	60	70	80	30	40	50	60	70	80	
-Bromotoluene	1134	1157	1179	1205	1244	1140	1152	1168	1183	1202	1223	–
-Bromotoluene	1135	1157	1180	1203	1240	1145	1156	1167	1181	1197	1211	–
-Bromotoluene	1132	1154	1176	1199	1235	1145	1154	1166	1178	1194	1208	–
-Chlorotoluene	1107	1129	1150	1174	1207	1115	1126	1138	1151	1168	1182	3.42
-Chlorotoluene	–	1127	1149	1172	1202	1117	1126	1137	1149	1162	1171	3.28
-Chlorotoluene	1096	1130	1144	1166	1192	1117	1128	1136	1146	1159	1168	3.33
,2-Dimethylbenzene	1071	1103	1130	1159	1199	1085	1102	1115	1125	1141	1150	3.12
,3-Dimethylbenzene	1088	1118	1147	1179	1215	1104	1120	1134	1147	1163	1174	3.20
,4-Dimethylbenzene	1091	1121	1155	1180	1219	1104	1122	1137	1150	1168	1176	3.15
-Methylacetophenone	894	896	895	898	902	888	887	886	889	892	892	–
-Methylacetophenone	900	900	899	900	903	887	884	881	882	886	887	–
-Methylacetophenone	898	900	897	896	899	879	875	873	874	879	882	2.10
-Methylanisole	1003	1026	1044	1063	1090	1014	1024	1030	1037	1043	1046	2.74
-Methylanisole	974	996	1011	1030	1055	985	992	996	1001	1003	1002	2.66
Methyl 2-methylbenzoate	984	991	997	999	1019	976	974	976	977	985	985	2.75
Methyl 3-methylbenzoate	1000	1003	1008	1008	1025	984	980	980	984	991	991	–
Methyl 4-methylbenzoate	1000	1001	1007	1009	1025	981	977	979	980	990	993	–
-Nitrotoluene	922	933	938	942	945	952	950	943	935	932	905	2.30
-Nitrotoluene	942	955	969	976	985	971	976	967	960	959	937	2.45
-Nitrotoluene	934	948	957	963	977	955	961	957	950	949	926	2.42
-Phenyltoluene	–	1296	1303	1318	1340	–	1274	1273	1274	1276	1277	–
-Phenyltoluene	–	1322	1336	1352	1379	–	1286	1286	1287	1289	1291	–
-Phenyltoluene	–	1334	1351	1367	1397	–	1294	1295	1297	1300	1304	–
-Tolualdehyde	869	873	877	881	894	868	871	870	880	893	879	2.26
-Tolualdehyde	872	873	876	876	890	873	877	876	884	901	877	–
-Tolualdehyde	863	865	866	868	883	861	863	862	873	890	872	–
2-Toluamide	628	655	617	609	589	621	584	561	550	568	631	–
3-Toluamide	688	682	668	654	638	650	619	594	581	597	625	1.18
4-Toluamide	689	677	664	651	639	664	633	592	582	594	642	1.18
2-Toluidine	728	747	747	742	730	765	766	764	762	769	761	1.32
3-Toluidine	732	751	748	739	731	771	769	763	761	762	754	1.41
4-Toluidine	740	755	750	745	722	770	767	763	761	770	760	1.39
2-Toluonitrile	871	872	870	863	860	898	895	889	884	886	874	–
3-Toluonitrile	882	884	884	873	869	914	912	906	900	905	880	–
4-Toluonitrile	877	878	878	865	857	906	903	897	890	891	871	–

[a] Values from refs. 11 and 15.

TABLE IV

RETENTION INDICES OF SUBSTITUTED PHENOLS

Compound	Retetion index (I)											Log P^a
	Methanol (%)					Acetonitrile (%)						
	40	50	60	70	80	30	40	50	60	70	80	
2-Aminophenol	576	592	580	562	569	586	589	600	618	635	711	0.52
3-Aminophenol	483	477	492	(467)	(474)	519	522	496	474	468	519	0.15
4-Aminophenol	446	446	(449)	(429)	515	463	471	445	423	424	516	0.04
2-Bromophenol	813	810	809	777	734	821	805	786	765	749	723	2.35
3-Bromophenol	862	860	849	830	793	859	838	815	792	772	743	2.63
4-Bromophenol	855	852	847	829	799	853	831	807	785	765	747	2.65
2-Chlorophenol	786	783	772	753	708	786	788	764	744	724	714	2.19
3-Chlorophenol	834	833	823	803	767	836	815	792	769	749	731	2.48
4-Chlorophenol	827	822	815	798	768	825	804	783	758	739	717	2.40
1,2-Dihydroxybenzene	599	589	583	570	568	577	587	565	558	553	687	1.01
1,3-Dihydroxybenzene	531	516	506	(483)	(476)	522	532	498	467	453	495	0.77
1,4-Dihydroxybenzene	469	474	461	(451)	(467)	467	492	461	433	435	493	0.50
2-Hydroxyacetophenone	853	860	867	875	888	849	867	853	854	855	857	1.90
3-Hydroxyacetophenone	679	661	641	619	614	663	669	616	605	588	597	1.39
4-Hydroxyacetophenone	649	624	594	551	(428)	615	606	569	562	556	562	1.30
2-Hydroxybenzaldehyde	782	783	783	769	771	862	875	834	822	845	867	1.81
3-Hydroxybenzaldehyde	656	641	623	600	571	648	630	610	592	571	(510)	1.38
4-Hydroxybenzaldehyde	600	565	507	(396)	(243)	600	577	550	537	513	(462)	1.35
2-Hydroxybenzamide	651	636	605	563	518	629	608	616	577	568	540	1.28
4-Hydroxybenzamide	463	454	(418)	(384)	(350)	420	393	(356)	(336)	(324)	(251)	0.33
2-Hydroxybenzonitrile	634	608	547	(447)	–	648	630	608	589	542	463	1.60
3-Hydroxybenzonitrile	693	678	658	625	510	711	688	665	639	610	606	1.70
4-Hydroxybenzonitrile	649	632	599	529	(196)	666	653	628	597	575	520	1.60
2-Methoxyphenol	715	711	705	698	682	720	712	697	693	682	678	1.32
3-Methoxyphenol	701	692	678	663	635	704	688	668	653	625	593	1.58
4-Methoxyphenol	666	656	643	628	606	668	654	636	622	601	565	1.39
Methyl 2-hydroxybenzoate	947	959	975	989	1006	931	944	940	945	941	940	2.55
Methyl 3-hydroxybenzoate	755	738	720	700	654	738	707	681	652	633	635	1.89
Methyl 4-hydroxybenzoate	744	733	701	666	565	708	688	667	649	622	608	1.96
2-Methylphenol	783	785	779	766	746	788	776	761	745	732	714	1.96
3-Methylphenol	775	768	764	748	725	772	758	741	722	708	691	1.96
4-Methylphenol	777	772	766	752	733	774	758	741	723	713	700	1.94
2-Nitrophenol	750	739	706	622	(228)	781	778	783	789	766	641	1.79
3-Nitrophenol	744	739	721	685	543	754	739	716	688	659	623	2.00
4-Nitrophenol	631	597	519	(379)	–	669	635	611	598	518	(395)	1.91
2-Phenylphenol	992	981	969	947	922	984	973	939	916	884	847	3.09
3-Phenylphenol	1007	989	973	943	918	985	950	914	888	854	822	3.23
4-Phenylphenol	1006	991	970	949	929	984	947	910	879	853	816	3.20

[a] Values from refs. 11 and 15.

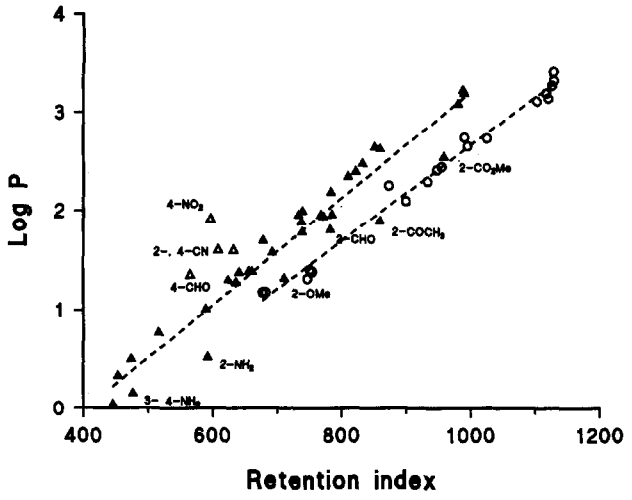


Fig. 1. Relationship of log P and retention indices for derivatives based on Tables III and IV. Analytes: ○ = substituted toluenes; ▲ = phenols; △ = suspected ionized phenols. Me = Methyl.

ABLE V

STIMATED RETENTION INDICES CALCULATED AS THE SUM OF PARENT INDEX AND SUB-ITUENT INDICES[a]

roups	I_{Sum}										
	Methanol (%)					Acetonitrile (%)					
	40	50	60	70	80	30	40	50	60	70	80
$+ I_{S,CH_3} + I_{S,Br}$	1128	1148	1168	1188	1210	1137	1154	1167	1178	1185	1190
$+ I_{S,CH_3} + I_{S,Cl}$	1099	1118	1136	1153	1171	1108	1125	1138	1149	1156	1161
$+ I_{S,CH_3} + I_{S,CH_3}$	1085	1113	1138	1161	1182	1110	1127	1140	1151	1158	1163
$+ I_{S,CH_3} + I_{S,COR} + I_{S,CH_3}$	900	900	900	900	900	900	900	900	900	900	900
$+ I_{S,CH_3} + I_{S,OR} + I_{S,CH_3}$	985	1001	1019	1036	1054	1000	1008	1012	1016	1015	1014
$+ I_{S,CH_3} + I_{S,CO_2R} + I_{S,CH_3}$	1000	1003	1006	1010	1015	990	993	994	996	996	996
$+ I_{S,CH_3} + I_{S,NO_2}$	950	959	966	971	976	969	973	971	964	952	936
$+ I_{S,CH_3} + I_{S,Ph}$	1306	1318	1332	1349	1368	1300	1298	1295	1295	1294	1295
$+ I_{S,CH_3} + I_{S,CHO}$	875	875	876	878	883	880	886	888	888	885	880
$+ I_{S,CH_3} + I_{S,CONH_2}$	704	692	679	666	653	679	636	613	614	636	681
$+ I_{S,CH_3} + I_{S,NH_2}$	749	759	760	755	741	795	797	796	796	795	793
$+ I_{S,CH_3} + I_{S,CN}$	876	879	877	871	861	913	918	914	910	901	890
$+ I_{S,OH} + I_{S,NH_2}$	443	430	402	363	310	484	454	427	405	385	370
$+ I_{S,OH} + I_{S,Br}$	822	819	810	796	779	825	811	798	787	776	767
$+ I_{S,OH} + I_{S,Cl}$	793	789	778	761	740	796	782	769	758	747	738
$+ I_{S,OH} + I_{S,OH}$	474	455	423	380	320	486	441	401	368	339	317
$+ I_{S,OH} + I_{S,COR} + I_{S,CH_3}$	594	571	542	508	469	588	557	530	508	492	477
$+ I_{S,OH} + I_{S,CHO}$	569	546	518	486	452	568	543	518	497	476	457
$+ I_{S,OH} + I_{S,CONH_2}$	398	363	321	274	222	367	293	244	223	227	258
$+ I_{S,OH} + I_{S,CN}$	570	550	519	479	430	601	575	545	519	492	467
$+ I_{S,OH} + I_{S,OR} + I_{S,CH_3}$	679	673	661	644	623	688	665	643	624	606	591
$+ I_{S,OH} + I_{S,CO_2R} + I_{S,CH_3}$	694	674	648	618	583	678	650	625	605	586	573
$+ I_{S,OH} + I_{S,CH_3}$	779	784	780	769	751	798	784	771	760	749	740
$+ I_{S,OH} + I_{S,NO_2}$	644	630	608	579	545	657	630	601	573	543	513
$+ I_{S,OH} + I_{S,Ph}$	1000	989	974	957	937	989	955	926	903	884	872

Values of I_P and $I_{S,X}$ derived from ref. 7.

TABLE VI

INTERACTION INCREMENTS FOR SUBSTITUTED TOLUENES

Substituent pairs	Interaction increment (δI)										
	Methanol (%)					Acetonitrile (%)					
	40	50	60	70	80	30	40	50	60	70	80
CH_3 + 2-Br	6	9	11	17	34	3	−2	1	5	17	33
CH_3 + 3-Br	7	9	12	15	30	8	2	0	3	14	21
CH_3 + 4-Br	4	6	8	11	25	8	0	−1	0	9	18
CH_3 + 2-Cl	8	11	14	21	36	7	1	0	2	12	21
CH_3 + 3-Cl	–	9	13	19	31	9	1	−1	0	6	10
CH_3 + 4-Cl	−3	12	8	13	21	9	3	−2	−3	3	7
CH_3 + 2-CH_3	−14	−10	−8	−2	17	−25	−25	−25	−26	−17	−13
CH_3 + 3-CH_3	3	5	9	18	33	−6	−7	−6	−4	5	11
CH_3 + 4-CH_3	6	8	17	19	37	−6	−5	−3	−1	10	13
CH_3 + 2-$COCH_3$	−6	−4	−5	−2	2	−12	−13	−14	−11	−8	−8
CH_3 + 3-$COCH_3$	0	0	−1	0	3	−13	−16	−19	−17	−14	−13
CH_3 + 4-$COCH_3$	−2	0	−3	−4	−1	−21	−25	−27	−26	−21	−18
CH_3 + 2-OCH_3	18	25	25	27	36	14	16	18	21	28	32
CH_3 + 3-OCH_3	−11	−5	−8	−6	1	−15	−16	−16	−15	−12	−12
CH_3 + 2-CO_2CH_3	−16	−12	−9	−11	4	−14	−19	−18	−19	−11	−11
CH_3 + 3-CO_2CH_3	0	0	−2	−2	10	−6	−13	−14	−12	−5	−5
CH_3 + 4-CO_2CH_3	0	−2	−1	−1	10	−9	−16	−15	−16	−6	−3
CH_3 + 2-NO_2	−28	−26	−28	−29	−31	−17	−23	−28	−29	−26	−31
CH_3 + 3-NO_2	−8	−4	3	5	9	2	3	−4	−4	7	1
CH_3 + 4-NO_2	−16	−11	−9	−14	1	−14	−12	−14	−14	−3	−10
CH_3 + 2-Ph	–	−22	−29	−31	−28	–	−24	−22	−21	−18	−18
CH_3 + 3-Ph	–	4	4	3	11	–	−12	−9	−8	−5	−4
CH_3 + 4-Ph	–	16	19	18	29	–	−4	0	2	6	9
CH_3 + 2-CHO	−6	−2	1	3	11	−12	−15	−18	−8	8	−1
CH_3 + 3-CHO	−3	−2	0	−2	9	−7	−9	−12	−4	16	−3
CH_3 + 4-CHO	−12	−10	−10	−10	0	−19	−23	−26	−15	5	−8
CH_3 + 2-$COHN_2$	−76	−33	−62	−57	−64	−58	−52	−53	−64	−68	−50
CH_3 + 3-$CONH_2$	−16	−10	−11	−12	−15	−29	−17	−20	−33	−39	−56
CH_3 + 4-$CONH_2$	−15	−15	−9	−15	−14	−15	−3	−22	−32	−42	−39
CH_3 + 2-NH_2	−21	−12	−12	−13	−11	−30	−31	−32	−34	−26	−32
CH_3 + 3-NH_2	−17	−8	−12	−14	−9	−24	−28	−33	−35	−33	−39
CH_3 + 4-NH_2	−9	−4	−10	−10	−19	−25	−30	−33	−35	−25	−33
CH_3 + 2-CN	−5	−7	−7	−8	−1	−15	−23	−26	−26	−15	−16
CH_3 + 3-CN	6	5	7	2	8	1	−6	−9	−10	4	−10
CH_3 + 4-CN	1	−1	1	−6	−4	−7	−15	−18	−20	−12	−19

.BLE VII

TERACTION INCREMENTS FOR SUBSTITUTED PHENOLS

bstituent irs	Interaction increment (δI)										
	Methanol (%)					Acetonitrile (%)					
	40	50	60	70	80	30	40	50	60	70	80
H + 2-NH_2	133	162	178	199	259	102	135	173	213	250	341
H + 3-NH_2	40	47	90	(104)	(164)	35	68	69	69	83	149
H + 4-NH_2	3	16	(47)	(66)	205	−21	17	18	18	39	146
H + 2-Br	−9	−9	−1	−19	−45	−4	−6	−12	−22	−27	−44
H + 3-Br	40	41	39	34	14	34	27	17	5	−6	−24
H + 4-Br	33	33	37	33	20	28	20	9	−2	−11	−20
H + 2-Cl	−7	−6	−6	−8	−32	−10	6	−5	−14	−23	−24
H + 3-Cl	41	44	45	42	27	40	33	23	11	12	−7
H + 4-Cl	34	33	37	37	28	29	22	14	0	−8	−21
H + 2-OH	125	134	160	190	248	91	146	164	190	214	370
H + 3-OH	57	61	83	(103)	(156)	36	91	97	99	114	178
H + 4-OH	−5	19	37	(71)	(147)	−19	51	60	65	96	176
H + 2-$COCH_3$	259	289	325	367	419	261	310	323	346	363	380
H + 3-$COCH_3$	85	90	99	111	144	75	112	86	97	96	120
H + 4-$COCH_3$	55	54	52	43	(−41)	27	49	39	54	64	88
H + 2-CHO	213	237	265	283	319	294	332	316	325	369	410
H + 3-CHO	87	95	105	114	119	80	87	92	95	95	(53)
H + 4-CHO[a]	31	19	−11	(−90)	(−209)	32	34	32	40	37	(5)
H + 2-$CONH_2$	253	273	284	289	296	262	315	372	354	341	282
H + 4-$CONH_2$	65	91	(97)	(110)	(128)	53	100	(112)	(113)	(97)	(−7)
H + 2-CN[a]	64	58	28	(−32)	–	47	55	63	70	50	−4
H + 3-CN	123	128	139	146	80	110	113	120	120	118	139
H + 4-CN[a]	79	82	80	50	(−234)	65	78	83	78	83	53
H + 2-OCH_3	36	38	44	54	59	32	47	54	69	76	87
H + 3-OCH_3	22	19	17	19	12	22	23	25	29	19	2
H + 4-OCH_3	−13	−17	−18	−16	−17	−14	−11	−7	−2	−5	−26
H + 2-CO_2CH_3	253	285	327	371	423	253	294	315	340	355	367
H + 3-CO_2CH_3	61	64	72	82	71	60	57	56	47	47	62
H + 4-CO_2CH_3	50	59	53	48	−18	30	38	42	44	36	35
H + 2-CH_3	4	1	−1	−3	−5	−10	−8	−10	−15	−17	−26
H + 3-CH_3	−4	−16	−16	−21	−26	−26	−26	−30	−38	−41	−49
H + 4-CH_3	−2	−12	−14	−17	−18	−24	−26	−30	−37	−36	−40
H + 2-NO_2[a]	106	109	98	43	(−317)	124	148	182	216	223	128
H + 3-NO_2	100	109	113	106	−2	97	109	115	115	116	110
H + 4-NO_2[a]	−13	−33	−89	(−200)	–	12	5	10	25	−25	(−118)
H + 2-Ph	−8	−8	−5	−10	−15	−5	18	13	13	−2	−25
H + 3-Ph	7	0	−1	−14	−19	−4	−5	−12	−15	−32	−50
H + 4-Ph	6	0	−4	−8	−8	−5	−8	−16	−24	−33	−56

Compounds considered to be partially ionized.

TABLE VIII

REGRESSION EQUATIONS RELATING CHANGE IN INTERACTION INCREMENT TO ELUENT CONCENTRATION FOR SUBSTITUTED TOLUENES

$I_{1,Y-Z} = ax^2 + bx + c$ (x = % modifier).

Substituent pairs	Coefficients of regression equation					
	Methanol			Acetonitrile		
	a	*b*	*c*	*a*	*b*	*c*
CH_3 + 2-Br CH_3 + 3-Br CH_3 + 4-Br CH_3 + 2-Cl CH_3 + 3-Cl CH_3 + 4-Cl	0.0150	−1.230	31	0.0209	−2.079	51
CH_3 + 2-CH_3	0.0243	−2.214	37			−22
CH_3 + 3-CH_3 CH_3 + 4-CH_3	0.0200	−1.670	40			0
CH_3 + 2-$COCH_3$			0			−11
CH_3 + 3-$COCH_3$ CH_3 + 4-$COCH_3$			0 0			−19
CH_3 + 2-OCH_3			26			21
CH_3 + 3-OCH_3			0			−14
CH_3 + 2-CO_2CH_3			−11			−15
CH_3 + 3-CO_2CH_3 CH_3 + 4-CO_2CH_3			0 0			−10
CH_3 + 2-NO_2			−28			−26
CH_3 + 3-NO_2			0			0
CH_3 + 4-NO_2			0			−11
CH_3 + 2-Ph			−28			−21
CH_3 + 3-Ph CH_3 + 4-Ph			0 21			0
CH_3 + 2-CHO CH_3 + 3-CHO CH_3 + 4-CHO			0 0 0	0.0060	−0.286	−17
CH_3 + 2-$CONH_2$	−0.0471	5.657	−219			−58
CH_3 + 3-$CONH_2$ CH_3 + 4-$CONH_2$			−13	−0.0148	0.972	−33
CH_3 + 2-NH_2 CH_3 + 3-NH_2 CH_3 + 4-NH_2			−12			−31
CH_3 + 2-CN			0			−20
CH_3 + 3-CN			0			0
CH_3 + 4-CN			0			−15

TABLE IX

REGRESSION EQUATIONS RELATING CHANGE IN INTERACTION INCREMENT TO ELUENT CONCENTRATION FOR INTERACTIONS OF SUBSTITUENTS WITH PHENOLIC HYDROXYL

Substituent pairs	Coefficients of regression equation					
	Methanol			Acetonitrile		
	a	*b*	*c*	*a*	*b*	*c*
OH + 2-NH_2	0.0479	−2.853	176	0.0510	−1.103	94
OH + 3-NH_2	0.0550	−3.550	93[a]	0.0387	−2.505	88
OH + 4-NH_2	0.1714	−16.031	378[a]	0.0759	−5.774	102
OH + 2-Br	−0.0557	5.868	−157	−0.0127	0.614	−11
OH + 2-Cl	−0.0371	3.937	−107	−0.0137	1.038	−23
OH + 3-Br OH + 4-Br OH + 3-Cl OH + 4-Cl			35	−0.0002	−0.968	63
OH + 2-OH	0.0728	−5.723	238	0.0945	−5.748	199
OH + 3-OH	0.0686	−5.828	181[a]	0.0144	0.640	19
OH + 4-OH	0.0857	−6.726	132[a]	0.0246	0.475	−36
OH + 2-CHO	0.0100	1.380	143	0.0455	−3.008	355
OH + 3-CHO	−0.0050	1.430	37			90
OH + 4-CHO[b]			–			35
OH + 2-$COCH_3$	0.0357	−0.305	214	−0.0257	5.048	138
OH + 3-$COCH_3$	0.0421	−3.667	166	0.0062	−0.150	85
OH + 4-$COCH_3$			50[c]	0.0161	−0.725	40
OH + 2-$CONH_2$	−0.0229	3.763	140	−0.1500	16.957	−114
OH + 4-$CONH_2$	−0.0064	2.221	−11[a]			–[d]
OH + 2-CN[b]			–	−0.0753	7.537	−118
OH + 3-CN	−0.1043	11.834	−191			120
OH + 4-CN[b]			–	−0.0384	4.080	−23
OH + 2-OCH_3	−0.0071	−0.237	33	−0.0036	1.470	−8
OH + 3-OCH_3			18	−0.0246	2.402	−30
OH + 4-OCH_3			−16	−0.0264	2.801	−77
OH + 2-CO_2CH_3	0.0300	0.660	178	−0.0302	5.542	176
OH + 3-CO_2CH_3			70			55
OH + 4-CO_2CH_3			53[e]			38
OH + 2-CH_3			0			−14
OH + 3-CH_3	0.0093	−1.544	43	−0.0064	0.227	−26
OH + 4-CH_3						−32
OH + 2-NO_2[b]			–	−0.1255	14.60	−217
OH + 3-NO_2			107			111
OH + 4-NO_2[b]			–			–
OH + 2-Ph			0	−0.0482	4.846	−105
OH + 3-Ph OH + 4-Ph	−0.0007	−0.606	28	−0.0204	1.317	−26

[a] Includes all data points including those based on capacity factors <0.2.
[b] Compound suspected of ionization particularly in methanol–buffer.
[c] Excludes increments based on capacity factors <0.2.
[d] Only two data points based on capacity factors >0.2.
[e] Mean excludes 80% methanol values as it appears to be an outlier.

composition were < 10 units, which is within the experimental errors of measurement [7], and they were assigned coefficients of zero. Changes in the increments across the eluent composition range of less than 20 units were regarded as insignificant and were assigned a single coefficient equal to the mean value. The halotoluenes were unusual as they all showed a systematic (although still relatively small) increase in interaction increment with the proportion of organic modifier. All six halo isomers were fitted to a common regression relationship. Similarly, the *meta* and *para* isomers of a number of the other substituents could be linked in a common equation.

The only significant toluene interactions were found for 2-methylbenzamide and in earlier work Clark *et al.* [20] reported that this compound was eluted more rapidly than the 3- and 4-isomers. They considered this difference to be due to a steric interaction causing the amide group to be less coplanar with the aromatic ring and hence more polar.

Most of the substituted phenols showed much larger interaction increments, which changed significantly with eluent composition (Table VII). The smallest effects were found for the methyl-, methoxy- and phenyl-substituted phenols, a number of which were assigned zero or constant regression coefficients (Table IX). Some of the *meta* and *para* substituents could again be linked in a common relationship. In developing the relationships for the other substituents, it was considered reasonable to include a number of the values which were possibly "unreliable" if these followed a steady trend. However, because there was a marked jump with 80% methanol for the 4-carbomethoxyl substituents, this value was thought to be an outlier and was excluded from the correlation.

The largest interactions were observed with the carbonyl substituents capable of hydrogen bonding, such as 2-hydroxyacetophenone (Fig. 2) and 2-hydroxybenzamide, which also differed markedly from the 4- and 3-isomers. These differences reflect those reported by Clark and co-workers [21,22] for the same or closely related compounds. Smaller but still significant interactions were also found for the *ortho*-dihydroxybenzenes and aminophenols.

When these quadratic expressions and constants for the toluenes and phenols were incorporated in the expert system program CRIPES, it was able to demonstrate

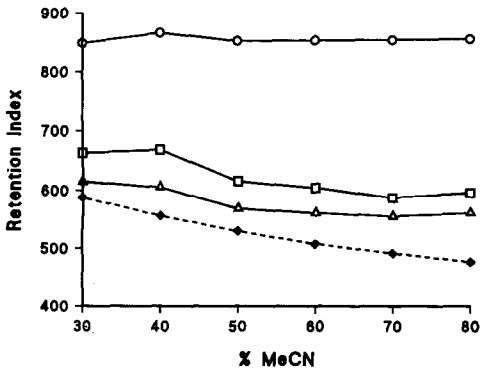


Fig. 2. Comparison of (◆) calculated (I_{Sum}) and experimental retention indices for (○) 2-, (□) 3- and (△) 4-hydroxyacetophenone. MeCN = Acetonitrile.

a reasonably successful ability to predict the retention indices and capacity factors of further substituted benzenes [10]. However, these interaction increments can really only be applied if values for the corresponding pairs of substituents are in the database. As a general prediction method this approach is very limited because a separate regression expression would have to be measured experimentally for each isomer of every possible pair of substituents. Additional terms would probably also be needed for multiple substitution. It was recognized, therefore, that the present form of the expert system database had only a limited application [10] and that a more general approach based on substituent susceptibilities would be needed.

General prediction model

For a more versatile prediction system, it is necessary to develop a model in which each substituent is associated with a set of terms that can reflect their mutual interactions in a similar manner to the σ and ρ terms used by Leo for the calculation of log P values [15,16]. Based on eqn. 4, an equation can be derived for $I_{I,X-Y}$:

$$I_{I,X-Y} = (\sigma_X\rho^*_Y + \sigma_Y\rho^*_X) + F^*_{HB} + F^*_o \quad (5)$$

where ρ^*, F^*_{HB} and F^*_o correspond to the terms in eqn. 4 but are expressed in retention index units. It is hoped that in each instance they could be directly related through a common regression equation for the eluent composition, *e.g.*, $\rho^* = \rho(ax^2 + bx + c)$, although the concept may need to be refined as more data becomes available. Leo [15] has noted that σ constants are valid for up to 80% organic modifier in aqueous solutions and so should be applicable in the present eluents. In preliminary calculations it appeared that the *meta* and *para* interactions differed so that instead of common σ values as suggested by Leo [15,16], published σ_{meta} and σ_{para} values [11] were used (Table X). The term for alkyl–aryl substitution has been omitted as it is thought that this effect may already be covered by an interaction term introduced earlier for alkyl substitution on a benzylic carbon [8].

Tsantili-Kakoulidou *et al.* [22] examined a very similar relationship between log k'_w values for a number of substituted phenols and anilines with sigma/rho correction values and *ortho* effects using linear regression analysis. Their results suggested that the approach was feasible but the weightings of the *ortho*-factors were quantized. They assigned ρ^* values but these often had large error ranges (*e.g.*, OH = 0.93 $\pm$ 0.30) and, unlike the Leo ρ values [15] some were negative (*e.g.*, Cl = -0.28 ± 0.21 and $NO_2 = -0.43 \pm 0.31$).

Meta and para groups

In order to determine the relationship between ρ and ρ^* in this study, the increments for the substituted phenols were examined. The phenolic group is a responder group ($\rho = 1.06$, σ relatively small) so that if σ is assumed to be zero (by analogy with Leo [15]), eqn. 5 for *meta* and *para* substituents can be redefined as

$$I_I = \sigma_X\rho^*_{OH} = \sigma_X\rho_{OH}(ax^2 + bx + c) \quad (6)$$

TABLE X

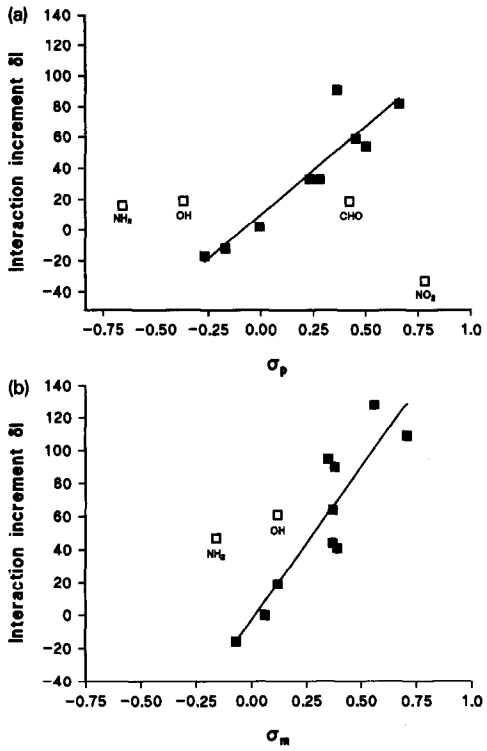
VALUES OF σ AND ρ USED IN CALCULATIONS OF INCREMENTS

Substituent	σ_{meta}[a]	σ_{para}[a]	ρ[b]
CH_3	−0.07	−0.17	0.00[c]
Phenyl	0.06	−0.01	0.00[c]
Inducers			
CN	0.56	0.66	0.00
NO_2	0.71	0.78	0.00
Br	0.39	0.23	0.00
Cl	0.37	0.23	0.00
Bi-directional			
CHO	0.35	0.42	0.44
CO_2CH_3	0.37	0.45	0.27
$COCH_3$	0.38	0.50	0.27
$CONH_2$	0.28	0.36	0.72
OCH_3	0.12	−0.27	0.50
Responders			
OH	0.12	−0.37	1.06
NH_2	−0.16	−0.66	1.08

[a] Ref. 11.
[b] Ref. 15.
[c] Ref. 14.

Thus in each eluent there should be a close relationship between the empirical interaction increments δI (from Table VII) and σ_X (from Table X). In methanol–buffer (50:50) a good linear correlation was found for the inducer and bi-directional substituents (*para*, Fig. 3a; *meta*, Fig. 3b). However, the amino and hydroxyl substituents, which are responder groups, (and the formyl and nitro substituents which gave ionized compounds) were clearly outliers. The σ values also correctly forecast the sign of the increments. Negative values of σ_{para} for methyl, methoxy and phenyl groups and of σ_{meta} for the methyl group were matched with negative retention increments and the positive value of σ_{meta} for the methoxyl group was matched by a positive increment (Table VII). Although σ_{meta} for the phenyl group gave the wrong indication, its influence was very small.

To determine the values of the coefficients a, b and c in eqn. 6, the ratios $\delta I_X/\rho_{OH}\sigma_X$ were then calculated for each substituent (using $\rho_{OH} = 1.06$ but excluding the hydroxyl and amino groups) and were correlated with the proportion of modifier (x). The phenyl and methyl groups were omitted as their σ values are very small and gave erratic ratios. The mean values of the ratios from the different substituents were virtually independent of the percentage of methanol (Fig. 4a for *para* substituents) and suggested that the relationship for methanol–buffer eluents could be represented by a single value rather than a quadratic expression, hence $\rho^*_{para\text{-}X} = 100\rho_X$ and $\rho^*_{meta\text{-}X} = 170\rho_X$. The results for most substituents in acetonitrile–buffer eluents were also similar so that $\rho^*_{para\text{-}X} = 105\rho_X$ and $\rho^*_{meta\text{-}X} = 190\rho_X$ (*e.g.*, for *meta* substituents see Fig. 4b). The exceptions were the bromo and chloro groups, whose ratios changed systematical-



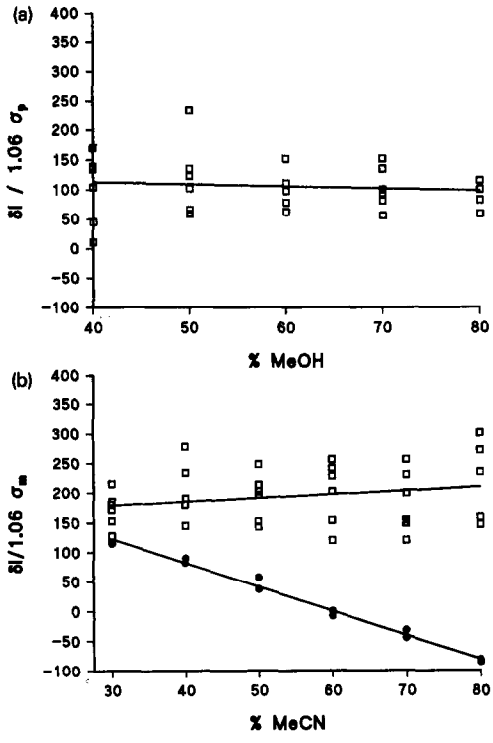


Fig. 3 (left). Relationship between interaction index (δI) and σ values in methanol–buffer (50:50). Open symbols were not used in the correlation. (a) *para*-Substituted phenols; (b) *meta*-substituted phenols.

Fig. 4 (right). Relationship between $\delta I_X/1.06\sigma_X$ and percentage of organic modifiers in eluents. (a) *para* Substituents in methanol (MeOH)–buffer; (b) *meta* substituents in acetonitrile (MeCN)–buffer, ● = bromo and chloro substituents.

ly with eluent composition so that $\rho^*_{para\text{-halogen}} = \rho_{halogen}(244-4x)$ and $\rho^*_{meta\text{-halogen}} = \rho_{halogen}(175-2.6x)$. A similar but smaller systematic change in the interaction increments also occurred with the halotoluenes (Table VIII).

Except for the halogens, these correction ratios suggest that for most substituents the interaction increment (I_1) can be assumed to be a constant irrespective of the proportion of modifier. This corresponds well to the empirical interaction increments (Table VII), many of which were almost constant across the eluent ranges for *para* and *meta* isomers. Using these ratios, the predicted interactions increments in methanol–buffer (60:40) and acetonitrile–buffer (60:40) were calculated using eqn. 5 (F^*_{HB} and $F^*_\sigma = 0$) and compared with the values from Table VII. In this case the reported values of σ for 3- and 4-hydroxyl were used with the bi-directional substituents [11]. In most instances the predicted I_1 and experimental values corresponded reasonably closely with differences of less than 30 units (Table XI). However, the values from hydroxyl and amino groups showed large and erratic errors and in their regression studies Fujita [14] found that these groups had to be regarded as

TABLE XI

PREDICTED INTERACTION INDEX VALUES FOR *PARA*- AND *META*-SUBSTITUTED PHENOLS

Substituent	Interaction terms[a]			
	Methanol		Acetonitrile	
	I_I	δI (60%)	I_I	δI (60%)
OH + 3-NH_2	−9	90	−32	69
OH + 4-NH_2	−110	(47)	−73	18
OH + 3-Br	70	39	8	5
OH + 4-Br	24	37	1	−2
OH + 3-Cl	67	45	7	11
OH + 4-Cl	24	37	1	0
OH + 3-OH	42	83	24	99
OH + 4-OH	−78	37	−41	65
OH + 3-$COCH_3$	74	99	82	97
OH + 4-$COCH_3$	43	52	46	54
OH + 3-CHO	72	105	80	95
OH + 4-CHO	28	−11[b]	30	40[b]
OH + 4-$CONH_2$	12	(97)	12	(113)
OH + 3-CN	101	139	113	120
OH + 4-CN	70	(80)[b]	73	78[b]
OH + 3-OCH_3	32	17	35	29
OH + 4-OCH_3	−47	−18	−49	−2
OH + 3-CO_2CH_3	73	72	80	47
OH + 4-CO_2CH_3	38	53	40	44
OH + 3-CH_3	−13	−16	−14	−38
OH + 4-CH_3	−18	−14	−19	−37
OH + 3-NO_2	128	113	143	115
OH + 4-NO_2	83	−89[b]	87	25
OH + 3-Ph	11	−1	12	−15
OH + 4-Ph	−1	−4	−1	−24

[a] I_I calculated using eqn. 5. δI from Table VII.
[b] Suspected ionized analyte.

outliers. Further studies will be needed to determine the best way to approach analytes containing two strong responder groups.

Using the same ratios for ρ^*/ρ it is also possible to predict the increments for the substituted toluenes using eqn. 5 (Table XII). In this instance ρ_Y^* for the methyl group is zero so that $I_I = \rho^*\sigma_Y$. Again, the predicted increments are constants irrespective of eluent composition and show a good correlation with the empirical values in methanol–buffer (60:40) or acetonitrile–buffer (60:40). In particular they reflect the high values for the amido substituent.

TABLE XII

PREDICTED INTERACTION INDEX VALUES FOR *PARA*- AND *META*-SUBSTITUTED TOLUENES

Substituent	Interaction terms[a]			
	Methanol		Acetonitrile	
	I_I	δI (60%)	I_I	δI (60%)
CH_3 + 3-Br	0	12	0	3
CH_3 + 4-Br	0	8	0	1
CH_3 + 3-Cl	0	13	0	0
CH_3 + 4-Cl	0	10	0	−3
CH_3 + 3-CH_3	0	9	0	−6
CH_3 + 4-CH_3	0	17	0	−3
CH_3 + 3-$COCH_3$	−3	−1	−4	−17
CH_3 + 4-$COCH_3$	−5	−3	−5	−26
CH_3 + 3-OCH_3	−6	−8	−6	−15
CH_3 + 4-OCH_3	−8	–	−1	–
CH_3 + 3-CO_2CH_3	−3	−2	−4	−12
CH_3 + 4-CO_2CH_3	−5	−1	−5	−16
CH_3 + 3-NO_2	0	3	0	−4
CH_3 + 4-NO_2	0	−9	0	−14
CH_3 + 3-Ph	0	4	0	−8
CH_3 + 4-Ph	0	19	0	2
CH_3 + 3-CHO	−5	0	−6	−4
CH_3 + 4-CHO	−7	−10	−8	−15
CH_3 + 3-$CONH_2$	−8	−11	−9	−33
CH_3 + 4-$CONH_2$	−13	−9	−13	−32
CH_3 + 3-NH_2	−13	−12	−14	−35
CH_3 + 4-NH_2	−18	−10	−13	−35
CH_3 + 3-CN	0	7	0	−10
CH_3 + 4-CN	0	1	0	−20

[a] I_I calculated using eqn. 5. δI values from Table VI.

Ortho substituents

For the *ortho*-substituents it is assumed that the σ/ρ electronic interactions are the same as those for the *para*-substituents and that $\sigma_{ortho} = \sigma_{para}$. Therefore, to determine the magnitude of any extra negative *ortho* interactions (F_o) the interaction increments of the *ortho*-substituents were compared with those for the *para*-substituents.

Only small differences (+10 to −30 units) were observed for most of the *ortho*-substituted toluenes (Table VI). Acetyl and cyano groups had a negligible *ortho* effect, bromo, chloro, formyl, amino and hydroxyl groups showed small positive effects and methyl, carbomethoxy and nitro groups showed small negative effects.

Only the bulky phenyl and amido groups (−21 to −61 units) showed a significant change which agreed with Leo's observation [15] that the *ortho* effect was greatest with $CONH_2$ > halogen > NO_2 > OH, NH_2, but his rankings contained anomalies and differed depending on the parent substituent.

Examination of those substituted phenols which do not undergo hydrogen bonding (Table VII) suggested that negative *ortho* effects are present for the bromo and chloro groups. These varied with eluent composition, rising sharply with increasing proportion of methanol but less markedly with acetonitrile. The phenyl group gave a small negative effect in methanol but a positive effect in acetonitrile and the methyl group was positive in both eluents. The assignment of negative *ortho* interaction indices for retention prediction will be difficult because insufficient examples are available in the present study for a detailed analysis. Leo [15] suggested a number of quantized assignments but many of his values were interpolations and he suggested that Taft steric effect constant, E_s, and field effects, F, might play a role.

The largest *ortho* interaction increments for the phenols were apparently due to hydrogen bonding and the substituents could be divided into three groups, weak interactions from methoxyl and possibly from nitro substituents, medium interactions with hydroxyl and amino groups and strong interactions with the carbonyl-containing substituents. In each instance the effect was very dependent on the eluent composition and could change by over 150 units. Leo [15] suggested that for a limited set of compounds a single F_{HB} factor could be used, although for some combinations an additional term was required. A more limited interaction (150–250 units) occurs between OH + OH and OH + NH_2 groups, which was very dependent on the eluent composition. To isolate the F_{HB} effect the σ/ρ increment (from Table XI) should first be subtracted from the interaction increments; however, as seen above, the predicted values for the hydroxyl and amino substituents are unreliable.

Until further examples have been studied it seems that it will be difficult to develop general rules for these pairs of substituents and they are best described by the empirical relationships determined earlier (Table IX). Because the structural features causing these interactions are well defined, in any prediction system it will be possible to make specific rules to accounts for these effects.

The interaction between nitro and hydroxyl groups is unclear. Leo suggested that no hydrogen-bonding correction was needed for 2-nitrophenol, but in the present study the changes were large and because of possible ionization direct comparison with the 4-nitro isomer is difficult. The predicted *para* interaction (Table IX) in methanolic eluents of 83 is close to the empirical *ortho* interaction in methanol–buffer (60:40) of 98, but deviates more in acetonitrile–buffer (60:40) with a predicted value of $I_1 = 87$ compared with the observed value $\delta I = 216$. Further studies at a lower pH will be needed to avoid ionization effects.

CONCLUSIONS

The model compounds show a number of interactions between substituents on an aromatic ring. These have been incorporated as empirical relationships into a retention scheme. It appears that a more widely applicable model can be based on the use of the sigma/rho corrections for *meta* and *para* interactions but that *ortho* effects, such as hydrogen bonding and steric effects, may need to be incorporated as specific

interaction terms for individual pairs of substituents. There were particular problems with compounds that could be ionized. Further model compounds will need to be examined to test the general model and specifically a method is needed to deal with analytes containing more than one electron-releasing group.

REFERENCES

1 J. L. Glajch and L. R. Snyder (Editors), *Computer-Assisted Method Development for High-Performance Liquid Chromatography*, Elsevier, Amsterdam, 1990.
2 M. De Smet, G. Musch, A. Peeters, L. Buydens and D. L. Massart, *J. Chromatogr.*, 485 (1990) 237.
3 R. Hindriks, F. Maris, F. Vink, A. Peeters, M. De Smet, D. L. Massart and L. Buydens, *J. Chromatogr.*, 485 (1990) 255.
4 Y. Baba, *J. Chromatogr.*, 485 (1990) 143.
5 T. Sasagawa, Y. Sakamoto, T. Hirose, T. Yoshida, Y. Kobayashi and Y. Sato, *J. Chromatogr.*, 485 (1990) 533.
6 P. Härmälä, H. Vuorela, P. Lehtonen and R. Hiltunen, *J. Chromatogr.*, 507 (1990) 367.
7 R. M. Smith and C. M. Burr, *J. Chromatogr.*, 475 (1989) 57.
8 R. M. Smith and C. M. Burr, *J. Chromatogr.*, 481 (1989) 71.
9 R. M. Smith and C. M. Burr, *J. Chromatogr.*, 481 (1989) 85.
10 R. M. Smith and C. M. Burr, *J. Chromatogr.*, 485 (1989) 325.
11 C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley-Interscience, New York, 1979.
12 T. Fujita, J. Isawa and C. Hansch, *J. Am. Chem. Soc.*, 86 (1964) 5177.
13 T. Fujita, *J. Pharm. Sci.*, 72 (1983) 285.
14 T. Fujita, *Prog. Phys. Org. Chem.*, 14 (1983) 75.
15 A. Leo, *J. Chem. Soc., Perkin Trans. 2*, (1983) 825.
16 A. J. Leo, in W. J. Dunn, J. H. Block and R. S. Pearlman (Editors), *Partition Coefficient. Determination and Estimation*, Pergamon Press, New York, 1986, p. 61.
17 E. P. Serjeant and B. Dempsey, *Ionisation Constants of Organic Acids in Aqueous Solutions*, Pergamon Press, Oxford, 1979.
18 T. Braumann, *J. Chromatogr.*, 373 (1986) 191.
19 K. Miyake, N. Mizuno and H. Terada, *J. Chromatogr.*, 439 (1988) 227.
20 C. R. Clark, M. J. M. Wells, R. T. Sansom, J. L. Humerick, W. B. Brown and B. J. Commander, *J. Chromatogr. Sci.*, 22 (1984) 75.
21 M. R. Clark, L. E. Garcia-Roura and C. R. Clark, *J. Liq. Chromatogr.*, 11 (1988) 3213.
22 C. R. Clark and L. E. Garcia-Roura, *J. Chromatogr. Sci.*, 27 (1989) 11.
23 A. Tsantili-Kakoulidou, N. El Tayar, H. van de Waterbeemd and B. Testa, *J. Chromatogr.*, 389 (1987) 33.