Journal of Chromatography, 475 (1989) 57–74
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 21 528

RETENTION PREDICTION OF ANALYTES IN REVERSED-PHASE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY BASED ON MOLECULAR STRUCTURE

I. MONOSUBSTITUTED AROMATIC COMPOUNDS

ROGER M. SMITH* and CHRISTINA M. BURR
Department of Chemistry, Loughborough University of Technology, Loughborough, Leics. LE11 3TU (U.K.)
(First received September 28th, 1988; revised manuscript received February 13th, 1989)

SUMMARY

A system has been developed to predict retentions in reversed-phase high-performance liquid chromatography based on the molecular structure of the analyte. The retentions are calculated as retention indices, on the alkyl aryl ketone scale, by the summation of a value for a parent compound, increments for the individual substituents and contributions for interactions between the substituents. In this paper the coefficients of the quadratic equations are reported, which define the increments for a range of 17 substituents on an aromatic ring over the eluent ranges 40–80% methanol in buffer and 30–80% acetonitrile in buffer.

INTRODUCTION

In many laboratories the development of a high-performance liquid chromatographic (HPLC) method for a new analyte is a process of trial and error coupled with experience of similar or closely related compounds. The first stage is the selection of a suitable eluent composition to give retentions within a reasonable time span. This has led to computer based prediction techniques, which can rapidly suggest suitable conditions for an isocratic elution based on an initial gradient elution[1]. In addition if more than one analyte is involved, the conditions for a suitable resolution can take a considerable time to achieve. As a consequence, there has been considerable interest in recent years in methods to aid the development of a separation according to predetermined resolution criteria[2,3]. Both these approaches combine experimental observation with calculation but in neither case is the structure of the analyte taken into account. Although these approaches have advantages for unknown compounds or impurities, the structure of most analytes is known. It should therefore be possible to make use of a knowledge of the elution strength and selectivity of the eluent towards different structural features to predict a potentially suitable eluent.

The primary aim of the present study has been to develop a method which can

predict the retention of a compound, based on its structure by the summation of contributions from the carbon skeleton and any substituents and thus suggest a suitable initial eluent for examination. Potentially this approach could also suggest the optimum conditions to achieve a particular resolution between two analytes but because the interactions between substituents are not fully understood particularly on heterocyclic and aromatic ring systems, it will probably not be possible to make accurate predictions for complex molecules. However, if the experimental retentions of a core molecule are known then it should be possible to predict the relative retentions of closely related compounds containing different substituents. Any deviations found in subsequent experiments could be used to examine the interactions between the functional groups.

The concept that individual substituents contribute to retention in a definable way has been studied in a number of laboratories. However, only in a few previous studies have attempts been made to use these relationships to predict retentions and in most of these cases only a limited range of substituents have been included.

Jandera[4–8] recorded values for the polar and non-polar contribution of different groups to retention and has discussed a method of predicting retention based on interaction indices[9]. In initial studies he described changes in test compounds with eluent composition and showed that the expressions could be used to predict the retention of these compounds. He compared the retentions with the *n*-alkylbenzenes which were considered to only possess significant "non-polar" interactions and each substituent was then identified by two parameters, a non-polar contribution (n_{ce}) and a specific or polar contribution (q_i). The n_{ce} values were found to be dependent on the parent compound while the q_i contributions were reported to be virtually independent of the nature of the rest of the molecule[7]. The resulting values were used to successfully predict capacity factors for a range of substituted benzenes and phenylurea and triazine herbicides.

There has also been a wide range of studies which have examined the effect of different substituents of the retention of analytes. These quantitative structure–retention relationships (QSRR) have been discussed in detail by Kaliszan[10]. Similar concepts have long been used as the basis of log partition coefficient (log P) calculations in quantitative structure–activity relationships (QSAR) studies. Hansch and Leo[11] have successfully shown that octanol–water partition coefficients can be calculated in an additive manner from the value of a parent compound, plus contributions for each substituent (π) and a similar approach by Rekker[12] has used fragmental constants (f factors). There is often a good correlation between the octanol–water partition coefficient and retention, particularly for structurally related compounds. However, this relationship is relatively poor if compounds containing different functional groups are compared. Kaliszan[10] listed over 100 studies relating log P to retention in either thin-layer chromatography (TLC) or HPLC. These studies include work by Hanai and co-workers[13–19] who have used the linear relationship between log P calculated using Rekker f constants and log k' to predict the relative retention of several different types of compounds including bases[13], phenols[14,15] and, in combination with dissociation constants, to predict the retention times of acids[16–19]. For each type of compound a separate regression equation was used to calculate the retention times of "unknown" members of a family of compounds. Octanol–water partition coefficients calculated using π values have also been used by Jinno and Kawasaki[20,21] to predict

the capacity factors of alkylbenzenes and polynuclear aromatic compounds. In a proposed general prediction method Jinno and co-workers[22–24] have used the π values as one of several descriptors which include molecular connectivity indices, number of electron donating/accepting groups and Hammett constants to calculate the capacity factors of a range of substituted benzenes. However, most of the studies of the relationship of retention and log P have been aimed at using HPLC to predict log P values rather than using log P to predict retentions.

One of the most commonly reported structure–retention relationships in HPLC is the linear relationship between carbon number and log k' and this forms the basis of most retention index scales proposed for HPLC[25] and be used to express the contribution of different groups to retention. Baker using a retention scale based on 2-ketoalkanes[26] showed a close linear relationship between the retention index values and log P of structurally related drug compounds[27]. Baker used this to develop a prediction method for the retention index of a compound from the measured index of a "parent" compound and a weighted value from the substituent Hansch substituent constant (π). The method was used to calculate retention indices of a number of drug compounds including barbiturates[27], anthranilic acid analogues[27], narcotic analgesic[28], steroids[29] and urushiols[30]. He also noted that the addition of groups such as the glucuronides to drug molecules caused predictable constant increments in the measured retention indices[31].

A small number of other workers have suggested the use of substituent contributions derived from retention indices to predict the retention of related compounds. Shalaby *et al.*[32] used the retention index scale to suggest a system to predict the retention indices of nitrogen bridged compounds based on measured log P values. Magg and Ballschmiter[33] derived functional group contributions for ergot alkaloids using the 2-ketoalkane scale and found that, although the retention indices of the compounds varied between columns, differences were not dependent on the column. However, the work was not extended to predict the retention indices of unknown compounds. Morishita *et al.*[34] have suggested a method of predicting retention indices, on an alkane scale, in which substituent contributions were calculated from monosubstituted benzenes. These were used in combination with terms to account for interactions between substituents to predict the retention indices of polysubstituted benzenes. This was a limited study and the group contributions were only determined at a single eluent composition. Popl and co-workers[35–37] have used a scale based on the number of aromatic rings to predict the retention of a range of "unknown" phenolic oxidants.

Rather than relate retention to log P or retention indices a number of workers have attempted to predict retentions from molecular structure by the use of a substituent or group contributions to capacity factors[38,39]. The definition of the "group effect" differs but is usually the difference between the retention of a substituted and unsubstituted compound and substituent contributions have been derived using many different "parent" species and experimental conditions. The compounds which have been examined include coumarins[40], catecholamines[41,42], 2-phenylethylamine derivatives[43], purines[39,44], chromonoid compounds[45], and substituted benzenes[46]. As the majority of the papers used substituent contributions based on log k' the contributions will depend on the separation conditions although it has been suggested that the derived values can be transferred from one ODS column to another[39].

Prediction method

The basis of the prediction system examined in the present 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 compound (*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 interactions). The retention index of a compound can then be determined as

$$RI = PI + SI_R + \Sigma SI_{Ar-X} + \Sigma SI_{R-X} + \Sigma II_{YZ}$$

where *PI* represents the retention index value of a parent compound; SI_R the retention index contribution from saturated aliphatic carbons; ΣSI_{Ar-X} the substituent index values for substituents on an aromatic ring; ΣSI_{R-X} the substituent index values for substituents on saturated aliphatic carbons (these will include olefin and carbonyl groups); and ΣII_{YZ} the interaction index values between substituents to account for H-bonding, and electronic effects.

The values of the retention indices and the increments will be dependent on the composition of the mobile phase. For most compounds it has been shown that there is a nearly linear relationship between percentage of composition and log k' but that a closer correlation can usually be obtained with a quadratic relationship particularly if a wide range of eluent compositions is being considered[47]. Consequently, for each different modifier, each of the terms in the prediction system will need to be defined as an experimentally determined quadratic equation of the form

$$I = ax^2 + bx + c$$

where x is the percentage of organic modifier in the eluent. It will be therefore possible to sum the a, b and c coefficients of the different components of the prediction equation to give an overall quadratic equation for each modifier

$$RI = \Sigma ax^2 + \Sigma bx + \Sigma c$$

Benzene was selected as the parent compound because all its substituted derivatives could be readily detected spectroscopically. A wide range of derivatives are also readily available, substituted both directly on the aromatic ring and on aliphatic side chains, which means that both types of substituents can be studied. In future work it is hoped to make the system more general so that other parent groups could be used.

It was decided to base the study on retention indices using the alkyl aryl ketone scale[48] rather than capacity factors or log capacity factors, because in previous studies in these laboratories, retention indices have been shown to be much more reproducible over time and are much less susceptible to small changes in the operating conditions (eluent composition, temperature and flow-rates) such as could occur between separations carried out on different occasions or on different equipment[49]. The retention indices of most compounds are also much less affected by the differences between brands of stationary phases than are capacity factors[50] and the intention was to develop a prediction system which would be generally applicable in other laboratories and if possible on other columns. The retention indices of the alkyl aryl

ketone standards are directly related to their capacity factors. Consequently, as long as the capacity factors of the alkyl aryl ketones are known on a particular column, the predicted retention indices of analytes can be converted to the corresponding estimated capacity factors. The alkyl aryl ketone scale has already been widely used in this laboratory for the study of the reproducibility of assay of drug compounds of forensic interest[25] and has been shown to be applicable to eluents containing methanol[48], acetonitrile[51], and tetrahydrofuran[51] and has been recently adopted in other laboratories as the basis of collections of retention values for drugs[52] and mycotoxins[53].

In this first part of the study retention parameters of benzene as the parent compound have been measured and the parameters, which describe the changes in retention due to the presence of single substituents on the aromatic ring have been determined. The robustness of the measurements, long-term precision studies, and the methods adopted to ensure reproducible results are described in the following paper[54]. In future papers the determination of expressions for substituents on aliphatic carbons, for isomers and for the interactions between groups in multisubstituted compounds will also be examined. A database of these expressions has been linked to an expert system (CRIPES, chromatographic retention index prediction expert system)[55] to provide a user friendly interface for the calculation of retention indices.

EXPERIMENTAL

Chemicals and eluents

Retention index alkyl aryl ketone standards (acetophenone, propiophenone, butyrophenone, valerophenone, hexanophenone and heptanophenone) and model aromatic compounds were purchased from various sources. Methanol and acetonitrile, HPLC grade, and sodium nitrate, A.R. grade, and disodium hydrogenorthophosphate and sodium dihydrogenorthophosphate, reagent grade, were from FSA Laboratory Supplies (Loughborough, U.K.).

Buffer solutions

Buffer solutions of pH 7 were prepared by adding disodium hydrogenorthophosphate (1.37 g) and sodium dihydrogenorthophosphate dihydrate (1.58 g) to 1000 ml of deionised water. For eluents containing 90% organic modifier the buffer was diluted ten-fold with water to avoid precipitation.

Sample solutions

Solutions of the retention index standards and model aromatic compounds were prepared in the mobile phases at a dilution which gave a signal at 254 nm using a 10-μl injection. The void volume marker was prepared as an aqueous solution containing 6 mg ml^{-1} of sodium nitrate.

HPLC equipment

HPLC separations were performed using a Pye-Unicam PU 4010 pump and a Pye-Unicam PU 4025 UV detector set at 254 nm. Injections of the samples (10 μl) were made using a Rheodyne 7125 valve fitted with a 20-μl loop. The column (100 × 5 mm I.D.) was packed with Spherisorb ODS-2, 5-μm (Batch 23/151, Phase Separations, Queensferry, U.K.). The column was maintained at a constant temperature

by enclosing it in a glass water jacket and circulating water at 30°C from a thermostated bath. Retention times were recorded on a Shimadzu Chromatopac C-R3A integrator.

Experimental procedure

Each set of separations consisted of the injection of solutions containing in turn, a mixture of the alkyl aryl ketones (acetophenone to heptanophenone), followed by three standard compounds (phenol, benzene and toluene), the individual model compounds and finally aqueous sodium nitrate (6 mg ml^{-1}) as a column void volume marker. This procedure was carried out in triplicate for each set of model compounds. Whenever possible the three runs were completed on a single day, however with eluents containing low organic modifier concentrations this was not practical even if higher flow-rates were used. The retention times were determined in a range of eluent compositions, methanol–buffer, pH 7.0 (40:60 to 90:10, v/v) and acetonitrile–buffer, pH 7.0 (30:70 to 90:10, v/v). Each set of data was collected using a single batch of eluent which was recycled.

Calculations of retention values

Capacity factors were calculated from the arithmetic mean of the triplicate retention times of each of the solutes using $k' = (t_R - t_0)/t_0$. To ensure consistency, the retention times (in min) were taken from the integrator to three decimal places and capacity factors were rounded to two decimal places. These were used to calculate retention indices (as integers), using the capacity factors of the alkyl aryl ketones included within the same set of injections as the model compounds. The least squares linear correlation between the log k' and the carbon number $\times$ 100 of the alkyl aryl ketones (acetophenone to heptanophenone) was determined as described previously[48] and the retention indices of the standard and model compounds were calculated by substitution into the regression equation.

RESULTS AND DISCUSSION

In order to determine the substituent contribution of different functional groups directly attached to an aromatic ring the capacity factors of the retention index scale alkyl aryl ketones (acetophenone to heptanophenone), the parent compound benzene and 16 mono-substituted aromatic model compounds were measured in a range of different eluent combinations of methanol–buffer, pH 7 (40:60 to 90:10) and acetonitrile–buffer, pH 7 (30:70 to 90:10) (Tables I and II). The monosubstituted model compounds covered a wide range of functional groups, however, it was not possible to examine carboxylic and sulphonic acids groups as they would be ionised in the buffer. Some of the groups could be considered as mixed alkyl aryl compounds (Ph–O–R, Ph–CO–R, Ph–CO_2–R and Ph–R) and in each case only the smallest homologue [R = methyl (Me)] was included in this study to derive a value for the aromatic functional group (Ph–X–). The effects of changes in the alkyl groups will be discussed in later papers.

The capacity factors were collected over a two year period and during this time the column had to be repacked a number of times with new stationary phase because the efficiency had deteriorated. Individual results may therefore have been obtained

TABLE I

CAPACITY FACTORS OF ALKYL ARYL KETONE STANDARDS AND MONOSUBSTITUTED MODEL COMPOUNDS IN ELUENTS CONTAINING METHANOL

Mobile phase, methanol–buffer, pH 7.

Compound	*Capacity factor*					
	Methanol (%)					
	40	*50*	*60*	*70*	*80*	*90*
Retention index standards						
Acetophenone	6.79	3.23	1.63	0.99	0.58	0.42
Propiophenone	15.74	6.61	2.93	1.58	0.83	0.52
Butyrophenone	34.56	12.92	5.01	2.40	1.12	0.62
Valerophenone	82.41	27.25	9.16	3.86	1.57	0.79
Hexanophenone	206.6	59.44	16.76	6.35	2.26	1.00
Heptanophenone	536.1	132.4	32.52	10.61	3.30	1.30
Monosubstituted model compounds						
Aniline	1.73	1.09	0.68	0.49	0.33	0.25
Anisole	12.23	6.08	3.33	1.85	1.05	–
Benzaldehyde	4.65	2.42	1.37	0.86	0.56	–
Benzamide	1.18	0.68	0.42	0.33	0.25	0.21
Benzene	11.49	7.37	3.58	2.00	1.08	0.64
Benzonitrile	5.01	2.83	1.37	0.86	0.50	0.35
Benzyl alcohol	2.31	1.31	0.79	0.56	0.41	0.24
Benzyl bromide	31.54	13.41	6.01	2.86	1.56	0.49
Benzyl chloride	24.56	10.90	5.04	2.48	1.24	0.49
Benzyl cyanide	4.79	2.28	1.20	0.70	0.48	0.26
Biphenyl	204.0	62.3	22.21	8.29	3.26	–
Bromobenzene	41.82	19.74	7.67	3.70	1.67	0.88
Chlorobenzene	32.52	15.79	6.44	3.16	1.46	0.79
Methyl benzoate	14.16	6.69	2.94	1.61	0.85	0.53
Nitrobenzene	8.12	4.67	2.32	1.35	0.74	0.46
Phenol	2.27	1.27	0.78	0.49	0.34	0.25
Toluene	29.57	13.66	6.81	3.38	1.68	0.86

on different columns under slightly different conditions and may not be directly comparable. No correction or standardisation was applied at this stage because this role will be provided by the conversion to retention indices. At high proportions of methanol or acetonitrile the capacity factors of many of the model compounds and standards are very small ($k' < 0.5$) so that their accuracy is likely to be sensitive to minor errors in the measurement of the retention times. Consequently these capacity factors may have a greater degree of uncertainty than results based on longer retention times.

The capacity factors for the alkyl aryl ketones, acetophenone to heptanophenone, in both sets of eluent combinations showed linear correlations between log k' and the retention index (carbon number $\times$ 100). The correlations were consistently good across the composition ranges (Table III, based on the corresponding capacity factors in Tables I and II) in agreement with earlier studies[48,51].

TABLE II

CAPACITY FACTORS OF ALKYL ARYL KETONE STANDARDS AND MONOSUBSTITUTED MODEL COMPOUNDS IN ELUENTS CONTAINING ACETONITRILE

Mobile phase, acetonitrile–buffer, pH 7.

Compound	*Capacity factor*						
	Acetonitrile (%)						
	30	*40*	*50*	*60*	*70*	*80*	*90*
Retention index standards							
Acetophenone	5.25	2.91	1.69	1.10	0.74	0.57	0.36
Propiophenone	11.94	5.71	2.89	1.71	1.07	0.75	0.45
Butyrophenone	25.06	10.31	4.57	2.49	1.47	0.96	0.55
Valerophenone	55.00	19.18	7.42	3.69	2.04	1.24	0.68
Hexanophenone	124.04	36.42	12.27	5.59	2.90	1.64	0.85
Heptanophenone	282.27	69.47	20.44	8.55	4.17	2.21	1.09
Monosubstituted model compounds							
Aniline	2.21	1.63	1.01	0.73	0.52	0.43	0.22
Anisole	13.43	6.85	3.43	1.98	1.17	0.81	–
Benzaldehyde	5.28	3.10	1.79	1.16	0.77	0.44	–
Benzamide	0.83	0.61	0.43	0.35	0.28	0.33	0.20
Benzene	12.52	6.57	3.42	2.04	1.27	0.87	0.45
Benzonitrile	5.86	3.27	1.82	1.15	0.74	0.55	0.28
Benzyl alcohol	1.76	1.15	0.80	0.57	0.43	0.39	0.32
Benzyl bromide	34.10	13.08	5.69	2.84	1.67	1.11	0.62
Benzyl chloride	27.63	11.05	4.97	2.53	1.50	0.94	0.42
Benzyl cyanide	6.88	3.50	1.93	1.13	0.75	0.53	0.36
Biphenyl	154.5	46.07	14.89	6.49	3.10	1.82	–
Bromobenzene	35.34	14.60	6.34	3.38	1.97	1.24	0.63
Chlorobenzene	28.58	12.32	5.51	3.01	1.77	1.13	0.57
Methyl benzoate	10.72	5.19	2.81	1.63	1.03	0.73	0.38
Nitrobenzene	9.08	4.70	2.43	1.45	0.89	0.62	0.30
Phenol	2.54	1.47	0.99	0.63	0.44	0.35	0.20
Toluene	30.63	11.95	6.29	3.02	1.86	1.23	0.58

Calculation of retention indices

The retention indices of the model compounds (Tables IV and V) were calculated from their capacity factors and those of the alkyl aryl ketones run as part of that same test set (usually on the same day). As expected, the retention indices of acetophenone (which has a defined retention index of 800), when calculated as a model monosubstituted compound (Ph–CO–Me) has experimental retention indices very close to this value with both organic modifiers. The small constant deviations in methanol eluents may suggest a small systematic error but this was not considered to be significant. To ensure consistency the defined value of $RI = 800$ has been used in subsequent stages in the calculations. Although retention indices of analytes are affected to only a small extent by small changes in the eluent conditions and composition[49], they do show some changes across relatively wide composition ranges. As expected the relative changes were considerably less than the corresponding changes

TABLE III

CORRELATION BETWEEN LOG k' AND CARBON NUMBER OF THE ALKYL ARYL KETONES

Log k' = a(100 × carbon number) + b.

	Correlation coefficient	*Slope* ($\times 10^3$)	*Intercept*
Methanol–buffer (v/v, %)			
40:60	0.9994	3.778	−2.132
50:50	0.9995	3.214	−2.080
60:40	0.9995	2.581	−1.865
70:30	0.9995	2.053	−1.658
80:20	0.9994	1.492	−1.434
90:10	0.9985	0.985	−1.179
Acetonitrile–buffer (v/v, %)			
30:70	0.9999	3.441	−2.033
40:60	0.9999	2.734	−1.718
50:50	0.9998	2.143	−1.481
60:40	0.9998	1.757	−1.359
70:30	0.9997	1.478	−1.308
80:20	0.9994	1.164	−1.179
90:10	0.9994	0.955	−1.213

TABLE IV

RETENTION INDICES OF MODEL COMPOUNDS IN METHANOL ELUENTS

Compound	*Retention index*					
	Methanol (%)					
	40	*50*	*60*	*70*	*80*	*90*
Acetophenone[a]	805	806	803	803	804	809
Aniline	650	658	657	659	639	579
Anisole	884	904	917	934	954	–
Benzaldehyde	774	777	777	775	784	–
Benzamide	605	589	578	570	551	514
Benzene	888	915	937	958	985	999
Benzene[b]	885	913	938	961	982	
Benzonitrile	776	788	775	774	760	736
Benzyl alcohol	689	691	698	684	675	610
Benzyl bromide	991	1004	1019	1030	1059	917
Benzyl chloride	962	976	992	992	994	915
Benzyl cyanide	773	766	763	738	722	658
Biphenyl	1205	1222	1231	1247	1270	–
Bromobenzene	1027	1051	1065	1088	1110	1139
Chlorobenzene	998	1021	1036	1051	1072	1090
Methyl benzoate	899	904	904	910	914	917
Nitrobenzene	851	857	864	874	874	853
Phenol	685	683	680	671	650	582
Toluene	987	1019	1039	1065	1095	1132

[a] Defined value 800.

[b] Parent index values for benzene values derived from quadratic regression equation (Table VI).

TABLE V

RETENTION INDICES OF MODEL COMPOUNDS IN ACETONITRILE ELUENTS

Compound	*Retention index*						
	Acetonitrile (%)						
	30	*40*	*50*	*60*	*70*	*80*	*90*
Acetophenone[a]	800	798	798	798	799	803	805
Aniline	691	706	694	695	691	696	656
Anisole	900	908	912	915	917	913	–
Benzaldehyde	781	786	786	788	788	779	–
Benzamide	568	549	521	511	509	593	636
Benzene	910	927	940	951	957	960	962
Benzene[b]	910	927	940	951	958	963	
Benzyl alcohol	654	640	630	624	636	645	690
Benzyl bromide	1026	1025	1020	1011	1002	986	985
Benzyl chloride	999	999	993	983	973	956	813
Benzyl cyanide	825	817	804	789	774	751	741
Benzonitrile	814	817	813	808	799	788	749
Biphenyl	1201	1198	1196	1195	1194	1195	–
Bromobenzene	1041	1054	1065	1074	1084	1093	1109
Chlorobenzene	1014	1027	1037	1045	1053	1058	1070
Methyl benzoate	890	890	900	894	894	897	894
Nitrobenzene	869	874	871	864	853	836	797
Phenol	695	687	674	660	639	645	671
Toluene	1005	1022	1036	1046	1054	1061	1072

[a] Defined value 800.

[b] Parent index values for benzene values derived from quadratic regression equation (Table VI).

in capacity factors. The effects were systematic up to 80% modifier but particularly for the acetonitrile separations, they were often non-linear (Fig. 1). The sharp changes in the retention indices for some compounds, such as the benzyl halides, aniline and phenol, in eluents containing 90% of organic modifier are apparently due to changes in the effective eluent conditions. The change to the lower buffer strength used with the highest proportions of organic modifier eluents should not be the cause of these effects as it has been found that the ionic strength has no effect on the retention indices of test compounds at 70% modifier[54]. Katz *et al.*[56] have suggested that there is a change in the active eluent composition at 90% methanol, which may result in a selectivity change in the system. As a consequence of these non-systematic changes it was decided to restrict the study to the composition ranges up to 80% and the results for the 90% proportion of modifiers have been omitted from the calculations. The measured values at 90% modifier are in any case rather uncertain because the corresponding retention times and capacity factors are so small.

Rather than use the retention indices for benzene in each eluent composition as the reference values for the database, it was decided to base the study on smoothed values (parent indices *PI*), calculated from the quadratic least squares relationship between the experimental retention index of benzene and the proportion of modifier in the eluent (up to 80% modifier, Table VI and Fig. 2).

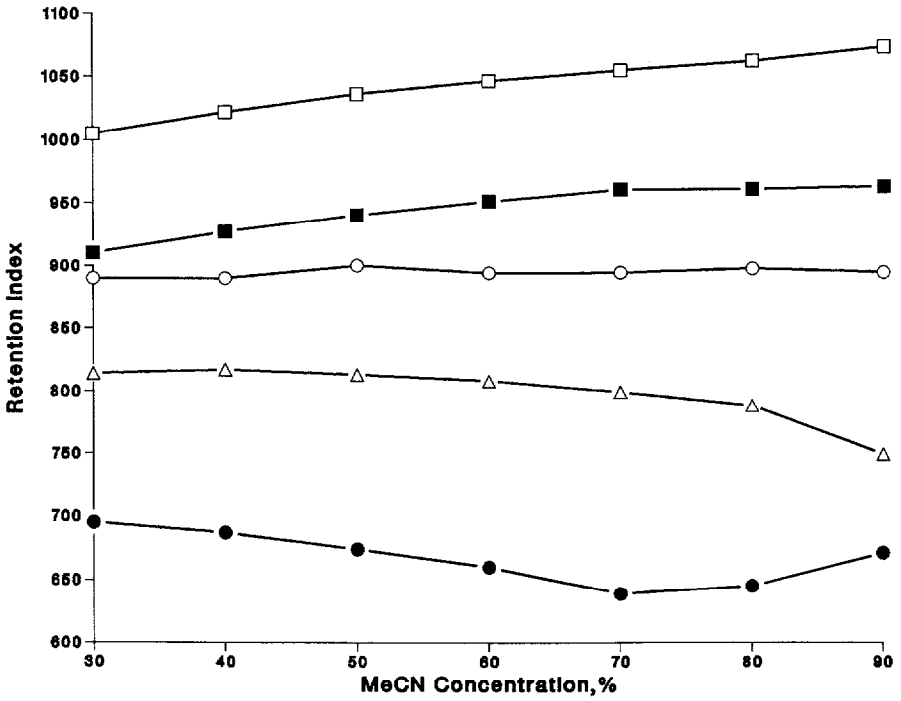

Fig. 1. Changes of retention indices of selected substituted aromatic compounds with percentage of acetonitrile (MeCN). Compounds: □, benzene; ■, toluene; ○, phenol; ●, methyl benzoate; △, benzonitrile.

Determination of substituent index equations

Using the calculated parent index (*PI*) values for benzene (given in Tables IV and V), the effects of each of the substituents can be calculated as the retention index increments (increment = $RI_{Ph-X} - PI_{Ph-H}$, Tables VII and VIII). Between 40 to 80% methanol and 30 to 80% acetonitrile the changes in the retention index increments for all the substituents are effectively systematic. The results for the methyl group were close to the defined value for the methylene increment of 100 units.

With each eluent the coefficients of the quadratic equation between the retention index increment and the percentage of composition were obtained (Table IX). If

TABLE VI

COEFFICIENTS OF PARENT INDEX EQUATIONS FOR BENZENE IN METHANOL AND ACETONITRILE CONTAINING ELUENTS

$PI = ax^2 + bx + c$. x = Percentage of modifier.

Organic modifier	*Range (%)*	*Coefficient*		
		a	*b*	*c*
Methanol	40–80	−0.0121	3.887	748
Acetonitrile	30–80	−0.0154	2.761	841

TABLE VII

RETENTION INDEX INCREMENTS FOR SUBSTITUENTS ON AN AROMATIC RING IN METHANOL ELUENTS

Increment = $RI_{\text{Model compound}} - PI$. PI = calculated parent index value for benzene (Table IV).

Substituent	*Hansch and Leo*[11] *π value*	*Retention index increment*				
		Methanol (%)				
		40	*50*	*60*	*70*	*80*
$CONH_2$	−1.49	−280	−324	−360	−391	−431
NH_2	−1.23	−235	−255	−281	−302	−343
CH_2OH	−1.03	−196	−222	−240	−277	−307
OH	−0.67	−205	−230	−258	−290	−332
CHO	−0.65	−111	−136	−161	−186	−198
CH_2CN	−0.57	−112	−147	−175	−223	−260
CN	−0.57	−109	−133	−163	−187	−222
$COCH_3$[a]	−0.55	−85	−113	−138	−161	−182
NO_2	−0.28	−34	−56	−74	−87	−108
OCH_3	−0.02	−1	−9	−20	−27	−28
CO_2CH_3	−0.01	14	−9	−34	−51	−68
H	0.00	0	0	0	0	0
CH_2Cl	0.17	77	63	54	31	12
CH_3[b]	0.56	102	106	101	104	113
Cl	0.71	113	108	98	90	90
CH_2Br	0.79	106	91	81	69	77
Br	0.86	142	138	127	127	128
Phenyl	1.96	320	309	293	286	288

[a] Based on defined values of RI = 800.
[b] Defined value = 100.

the variation in the increment across the composition range was less than 10 units a single mean value was included instead of a quadratic expression. Values for the aromatic substituent (Ph–X–) of the mixed alkyl aryl substituents (Ph–X–R), were calculated by excluding the contributions from the aliphatic group (–Me, SI = 100). These coefficients will be used in the prediction scheme to calculate the substituent index (SI) values as $SI = ax^2 + bx + c$.

Prediction of retention indices

The substituent index equation coefficients can now be used to calculate the predicted substituent indices and the retention indices for the model compounds. The correspondence between the experimental and the calculated substituent indices is close (Fig. 3) and should be sufficiently accurate for reliable prediction calculations. However, the calculated indices should not be extrapolated to compositions outside the measured regions as the values will be unreliable.

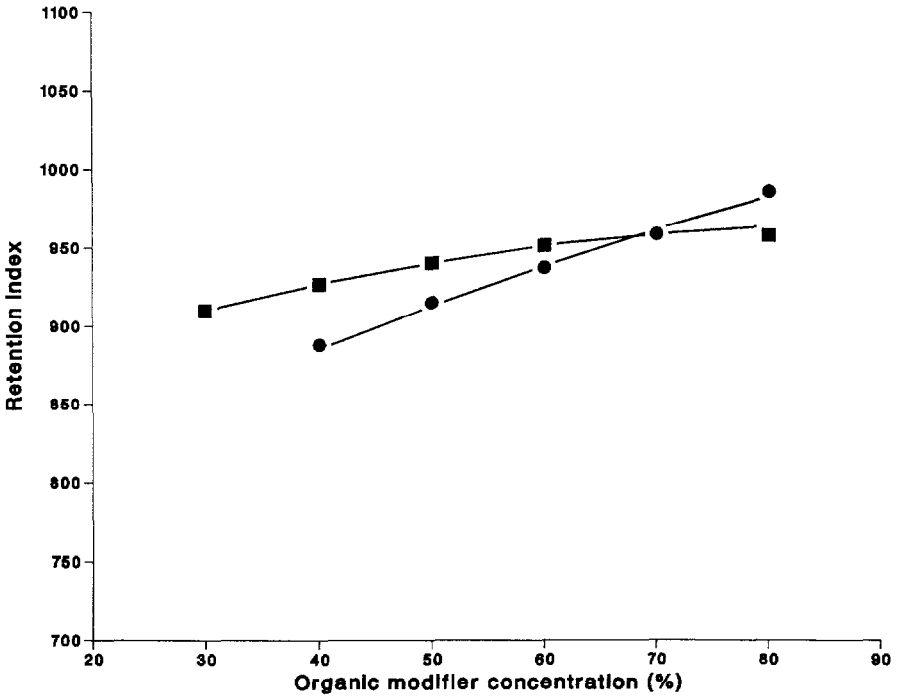

Fig. 2. Comparison of experimental retention indices of benzene in methanol–buffer and acetonitrile–buffer eluents with calculated values of parent index values derived from quadratic relationships (Table VI). Eluents: ●, methanol–buffer; ■, acetonitrile–buffer. Points are measured retention indices and curves are calculated values.

Relationship between substituent indices and octanol–water partition substituent increments

Within groups of closely related compounds, log k' of analytes in reversed-phase (RP)-HPLC has frequently been linearly correlated with the octanol–water partition coefficient (log P) (ref. 10) which are calculated in an additive manner from the Hansch substituent constants (π) and the octanol–water log P value of a parent. The values for the substituent indices should be thus related to the increments reported for the prediction of octanol–water partition coefficients, although the exact values will differ because of the different organic phases involved. This relationship could provide an mechanism by which estimated SI values could be obtained for substituents not determined experimentally (*e.g.*, the non-ionised carboxylic acid group). However, care must be taken as this approach would be expected give much less reliable values than the experimental data as the interactions occurring in an RP-HPLC separation are not always directly comparable to those in the octanol–water partition system.

The correlations between the π values for the aromatic substituents (listed in Table VII and VIII) and the corresponding SI values at different eluent compositions have therefore been determined (Table X). These results suggest that the relationship is approximately linear in both methanol and acetonitrile containing eluents. A com-

TABLE VIII

RETENTION INDEX INCREMENTS FOR SUBSTITUENTS ON AN AROMATIC RING IN ACETONITRILE ELUENTS

Increment = $RI_{\text{Model compound}} - PI$. PI = calculated parent index value for benzene (Table V).

Substituent	*Hansch and Leo*[11] *π value*	*Retention index increment*					
		Acetonitrile (%)					
		30	*40*	*50*	*60*	*70*	*80*
$CONH_2$	−1.49	−342	−378	−419	−440	−449	−370
NH_2	−1.23	−219	−221	−246	−256	−267	−267
CH_2OH	−1.03	−256	−287	−310	−327	−322	−318
OH	−0.67	−215	−240	−266	−291	−319	−318
CHO	−0.65	−129	−141	−154	−163	−170	−184
CH_2CN	−0.57	−85	−110	−136	−162	−184	−212
CN	−0.57	−96	−110	−127	−143	−151	−175
$COCH_3$[a]	−0.55	−110	−127	−140	−151	−158	−163
NO_2	−0.28	−41	−53	−69	−87	−105	−127
OCH_3	−0.02	−10	−19	−28	−36	−41	−50
CO_2CH_3	−0.01	−20	−37	−40	−57	−64	−66
H	0	0	0	0	0	0	0
CH_2Cl	0.17	89	72	53	32	15	−7
CH_3[b]	0.56	95	95	96	94	95	95
Cl	0.71	104	100	97	94	95	95
CH_2Br	0.79	116	98	80	60	44	23
Br	0.86	131	127	125	123	126	130
Phenyl	1.96	291	271	256	244	236	232

[a] Based on defined values of RI = 800.
[b] Defined value = 100.

parison of the values for SI in methanol–buffer (40:60) and π values (Fig. 4) showed a good correlation, with only one major outlier which was identified as the phenolic hydroxyl group. The equivalent curves for acetonitrile–buffer (40:60) (Fig. 5) suggested three outliers, the phenolic hydroxyl, benzyl hydroxyl and the carboxamide groups, again all are hydrogen bonding species. Differences in the relationship between log P and log k' particularly between hydrogen bonding species and non-hydrogen bonding species have been noted previously[57]. The overall pattern for acetonitrile was more scattered suggesting that the octanol–water partition is a poorer model of the interactions on HPLC. The values are closely related to those used for the prediction of octanol–water partition constants.

CONCLUSIONS

The parent retention indices for benzene and substituent indices for 17 aromatic substituents have been determined and expressed as quadratic equations covering a

TABLE IX

COEFFICIENTS OF SUBSTITUENT INDEX EQUATIONS FOR SUBSTITUENTS ON AN AROMATIC RING

Methanol–buffer (40:60 to 80:20) and acetonitrile–buffer (30:70 to 80:20). $SI = ax^2 + bx + c$.

Substituent Ph–X	*Methanol–buffer*			*Acetonitrile–buffer*		
	a	*b*	*c*	*a*	*b*	*c*
$CONH_2$	0.0093	−4.804	−104	0.1260	−14.878	2
NH_2	−0.0264	0.541	−215	0.0118	−2.405	−153
CH_2OH	−0.0193	−0.456	−148	0.0513	−6.872	−95
OH	−0.0271	0.117	−167	0.0218	−4.616	−93
CHO	0.0186	−4.469	39	0.0025	−1.335	−92
CH_2CN	−0.0171	−1.663	−18	0.0002	−2.543	−9
CN	−0.0114	−1.429	−34	−0.0025	−1.251	−57
$COCH_3$	0.0114	−3.791	48	0.0150	−2.704	−43
CO-R[a]	0.0114	−3.791	−52	0.0150	−2.704	−143
NO_2	0.0050	−2.390	53	−0.0104	−0.586	−14
OCH_3	0.0129	−2.263	70	0.0029	−1.097	20
O-R[a]	0.0129	−2.263	−30	0.0029	−1.097	−80
CO_2CH_3	0.0143	−3.774	143	0.0105	−2.096	33
CO_2-R[a]	0.0143	−3.774	43	0.0105	−2.096	−67
H	0	0	0	0	0	0
CH_2Cl	−0.0171	0.437	86	−0.0030	−1.586	140
CH_3[b]	0	0	100	0	0	100
Cl	0.0086	−1.669	167	0	0	98
CH_2Br	0.0314	−4.571	240	−0.0012	−1.711	168
Br	0.0150	−2.190	207	0	0	127
Phenyl	0.0250	−3.870	436	0.0193	−3.299	372

[a] Values exclude the contribution from the saturated aliphatic R group (Me = 100).
[b] Defined value.

TABLE X

REGRESSION COEFFICIENTS FOR π COMPARED TO CALCULATED SUBSTITUENT INDICES

$SI = a\pi + b$.

	Coefficients		*Correlation coefficient*
	a	*b*	
Methanol (%)			
40	−7.97	178	0.9849
50	−26.62	185	0.9839
60	−44.52	195	0.9816
70	−62.00	207	0.9787
80	−70.58	222	0.9607
Acetonitrile (%)			
30	−18.95	180	0.9739
40	−35.32	189	0.9661
50	−55.54	194	0.9573
60	−58.91	199	0.9588
70	−70.04	197	0.9622
80	−77.64	194	0.9627

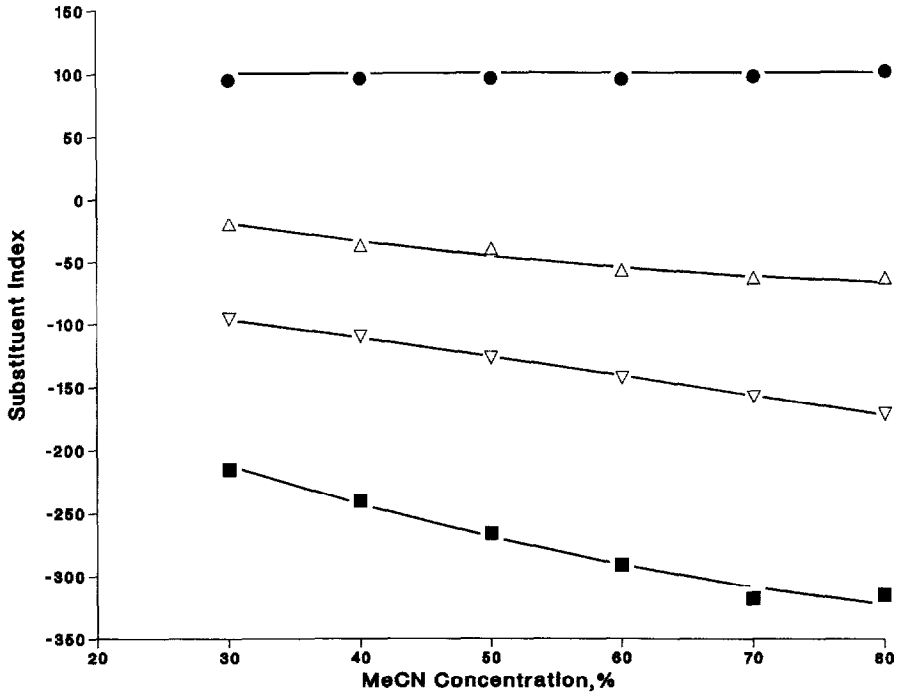

Fig. 3. Comparison of experimental values of retention index increments and calculated values of substituent indices in acetonitrile (MeCN)–buffer. Points are experimental values and curves are calculated substituent indices. Compounds: ■, phenol; ●, toluene; △, methyl benzoate; ▽, benzonitrile.

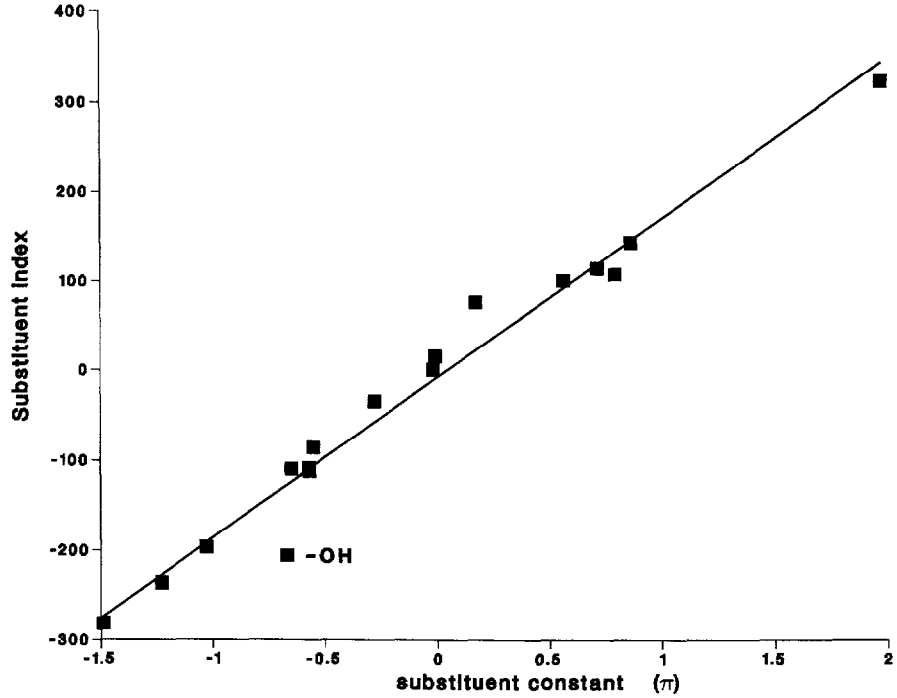

Fig. 4. Comparison of Hansch π values for substituents and calculated substituent indices in methanol–buffer (40:60).

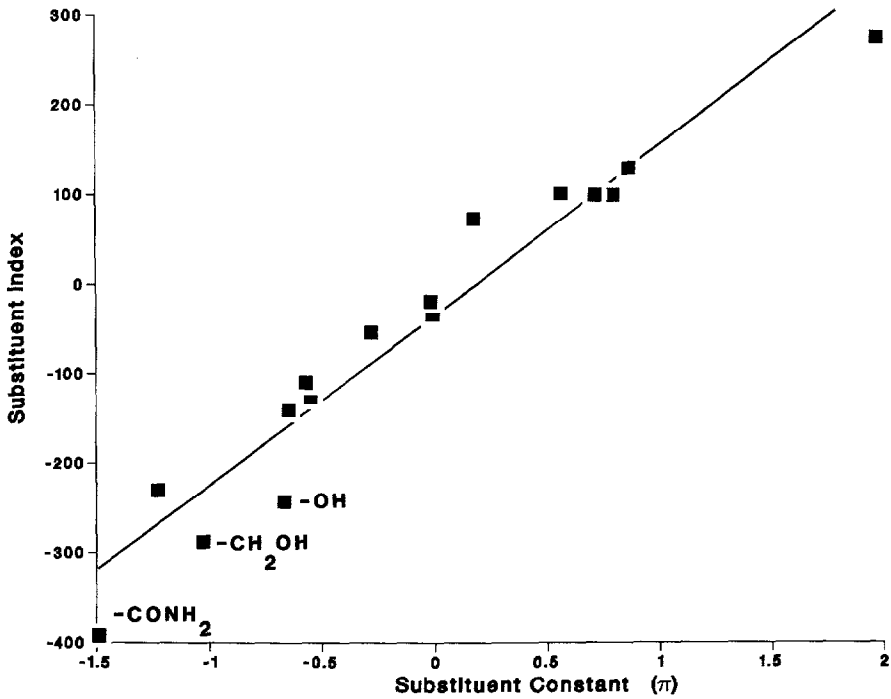

Fig. 5. Comparison of Hansch π values for substituents and calculated substituent indices in acetonitrile–buffer (40:60).

wide range of methanol and acetonitrile containing eluents, which can be used to predict retention in multisubstituted compounds.

ACKNOWLEDGEMENTS

Thanks are expressed to the Science and Engineering Research Council for a research grant and for a studentship to C.M.B. and to Phase Separations Ltd., U.K., for a gift of Spherisorb ODS-2.

REFERENCES

1 L. R. Snyder, J. L. Glajch and J. J. Kirkland, *Practical HPLC Method Development*, Wiley-Interscience, New York, 1988, pp. 239–245.
2 J. C. Berridge, *Techniques for the Automated Optimization of HPLC Separations*, Wiley, Chichester, 1985.
3 P. J. Schoenmakers, *Optimization of Chromatographic Selectivity (Journal of Chromatography Library*, Vol. 35), Elsevier, Amsterdam, 1986.
4 P. Jandera, *J. Chromatogr.*, 314 (1984) 13.
5 P. Jandera, *J. Chromatogr.*, 352 (1986) 91.
6 P. Jandera, *J. Chromatogr.*, 352 (1986) 111.
7 P. Jandera and M. Špaček, *J. Chromatogr.*, 366 (1986) 107.
8 P. Jandera, *Chromatographia*, 19 (1984) 101.
9 H. Colin, G. Guiochon and P. Jandera, *Anal. Chem.*, 55 (1983) 442.

10 R. Kaliszan, *Quantitative Structure–Chromatographic Retention Relationships* (*Chemical Analysis*, Vol. 93), Ed. J. D. Winefordner, Wiley, New York, 1987.
11 C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York, 1979.
12 R. F. Rekker, *The Hydrophobic Fragmental Constant*, Elsevier, Amsterdam, 1977.
13 T. Hanai and J. Hubert, *J. Liq. Chromatogr.*, 8 (1985) 2463.
14 T. Hanai and J. Hubert, *J. High Resolut. Chromatogr. Chromatogr. Commun.*, 6 (1983) 20.
15 T. Hanai, C. Tran and J. Hubert, *J. High Resolut. Chromatogr. Chromatogr. Commun.*, 4 (1981) 454.
16 T. Hanai, K. C. Tran and J. Hubert, *J. Chromatogr.*, 239 (1982) 385.
17 T. Hanai and J. Hubert, *J. Chromatogr.*, 239 (1982) 527.
18 T. Hanai and J. Hubert, *J. High Resolut. Chromatogr. Chromatogr. Commun.*, 7 (1984) 524.
19 T. Hanai, *J. Chromatogr.*, 332 (1985) 189.
20 K. Jinno and K. Kawasaki, *Anal. Chim. Acta*, 152 (1983) 25.
21 K. Jinno and K. Kawasaki, *Chromatographia*, 17 (1983) 337.
22 K. Jinno and M. Kuwajima, *Chromatographia*, 21 (1986) 541.
23 K. Jinno and K. Kawasaki, *J. Chromatogr.*, 316 (1984) 1.
24 K. Jinno and K. Kawasaki, *J. Chromatogr.*, 298 (1984) 326.
25 R. M. Smith, *Adv. Chromatogr.*, 26 (1987) 277.
26 J. K. Baker and C.-Y. Ma, *J. Chromatogr.*, 169 (1979) 107.
27 J. K. Baker, *Anal. Chem.*, 51 (1979) 1693.
28 J. K. Baker, R. E. Skelton, T. N. Riley and J. R. Bagley, *J. Chromatogr. Sci.*, 18 (1980) 153.
29 J. K. Baker and E. K. Fifer, *J. Pharm. Sci.*, 69 (1980) 590.
30 C.-Y. Ma, M. A. Elsohly and J. K. Baker, *J. Chromatogr.*, 200 (1980) 163.
31 J. K. Baker, *J. Liq. Chromatogr.*, 4 (1981) 271.
32 A. Shalaby, Zs. Budvári-Brány and Gy. Szász, *J. Liq. Chromatogr.*, 7 (1984) 1133.
33 H. Magg and K. Ballschmiter, *J. Chromatogr.*, 331 (1985) 245.
34 F. Morishita, H. Kakihana and T. Kojima, *Anal. Lett.*, 17 (1984) 2385.
35 M. Popl, V. Dolanský and J. Mostecký, *J. Chromatogr.*, 91 (1974) 649.
36 M. Popl, V. Dolanský and J. Mostecký, *J. Chromatogr.*, 117 (1976) 117.
37 M. Popl, I. Vít and F. Šmejkal, *J. Chromatogr.*, 213 (1981) 363.
38 V. Ya. Davydov, *J. Chromatogr.*, 365 (1986) 123.
39 S. P. Assenza and P. R. Brown, *J. Chromatogr.*, 282 (1983) 477.
40 K. Glowniak and M. L. Bieganowska, *J. Chromatogr.*, 370 (1986) 281.
41 I. Molnár and Cs. Horváth, *J. Chromatogr.*, 145 (1978) 371.
42 B.-K. Chen and Cs. Horváth, *J. Chromatogr.*, 171 (1979) 15.
43 R. Gill, S. P. Alexander and A. C. Moffat, *J. Chromatogr.*, 218 (1981) 639.
44 P. R. Brown and E. Grushka, *Anal. Chem.*, 52 (1980) 1210.
45 J. Borda, V. Szabó and J. Kelemen, *J. Chromatogr.*, 286 (1984) 113.
46 E. Tomlinson, H. Poppe and J. C. Kraak, *Int. J. Pharm.*, 7 (1981) 225.
47 P. J. Schoenmakers, H. A. H. Billiet and L. de Galan, *J. Chromatogr.*, 185 (1979) 179.
48 R. M. Smith, *J. Chromatogr.*, 236 (1982) 313.
49 R. M. Smith, T. G. Hurdley, R. Gill and A. C. Moffat, *Chromatographia*, 19 (1987) 401.
50 R. M. Smith, *Anal. Chem.*, 56 (1984) 256.
51 R. M. Smith, G. A. Murilla and C. M. Burr, *J. Chromatogr.*, 388 (1987) 37.
52 D. W. Hill and K. J. Langner, *J. Liq. Chromatogr.*, 10 (1987) 377.
53 J. C. Frisvad and U. Thrane, *J. Chromatogr.*, 404 (1987) 195.
54 R. M. Smith and C. M. Burr, *J. Chromatogr.*, 475 (1989) 75.
55 C. M. Burr and R. M. Smith, *Anal. Proc.*, 26 (1989) 26.
56 E. D. Katz, K. Ogan and R. P. W. Scott, *J. Chromatogr.*, 352 (1986) 67.
57 K. Miyake, N. Mizuno and H. Terada, *Chem. Pharm. Bull.*, 34 (1986) 4787.