CHROM. 23 437

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

VII^a. Separations using tetrahydrofuran-buffer eluents

ROGER M. SMITH* and RUI WANG

Department of Chemistry, Loughborough University of Technology, Loughborough, Leics. LE11 3 TU (UK) (First received March 1 lth, 1991; revised manuscript received May 6th, 1991)

ABSTRACT

As part of the development of a technique to predict retention in high-performance liquid chromatography from the molecular structure of analytes by the summation of contributions representing the parent skeleton and functional groups, the retention index values for benzene and the substituent indices for 21 aromatic and 4 aliphatic substituent groups, based on the alkyl aryl ketone scale, have been determined in 20:80 to 60:40 tetrahydrofuran-pH 7.0 buffer eluents on a Spherisorb ODS-2 column. The results have been compared with index values obtained with eluents based on methanol and acetonitrile.

INTRODUCTION

In recent years there has been considerable interest in the application of computer-based systems to predict retention times and relative retention times in highperformance liquid chromatography (HPLC) [1]. Previous studies from these laboratories have examined predictions based on the components of the molecular structure of the analyte [2-4]. To increase the applicability of the system the calculations were based on the alkyl aryl ketone retention index scale. The retention index (I) of the analyte is given by the summation of a series of terms:

$$
I = I_{P} + I_{S,R} + \sum I_{S,R-X} + \sum I_{S,Ar-X} + \sum I_{I,Y-Z}
$$
 (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 ($I_{S,R-X}$), contributions for aromatic substituents ($I_{S,Ax-X}$), and terms to account for any interactions between substituents $(I_{1,Y-Z})$ caused by electronic, hydro-

0021-9673/91/\$03.50 (C 1991 Elsevier Science Publishers B.V.

a For part VI, see ref. 4.

gen 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 using a quadratic equation $(I = ax^2 + bx + c$ in which $x = \frac{9}{6}$ organic modifier and a , b and c are index coefficients). So far in this study [3,4] the coefficients for benzene as a parent compound and for the substituent and interaction indices of a wide range of aromatic and aliphatic functional groups have been determined for separations carried out using methanol-pH 7 buffer and acetonitrile-pH 7 buffer eluents. The terms have been included in an expert systems program CRIPES ("chromatographic retention index prediction expert system") which has been tested with a range of substituted analytes [3].

In studies to improve or optimise the selectivity of a separation, tetrahydrofuran (THF) has frequently been proposed as a useful alternative organic modifier, instead of methanol or acetonitrile, with different interaction and retention properties [5,6]. It would therefore be expected that the index values of the parent compound and the different substituents should differ when determined in each of the three eluents. An understanding of these relative changes for compounds containing different groups should make possible the prediction of the optimum conditions for the separation of the components of a mixture.

In the present work the parent and substituent index coefficients for a number of aromatic and aliphatic functional groups are reported for separations carried out using THF-pH 7.0 buffer eluents over the range 20-60% THF and are compared with the indices determined earlier for methanol and acetonitrile-buffer eluents. The potential application of the alkyl aryi ketone retention index scale in THF-water eluents has been demonstrated in a preliminary study [7], which examined the effect of different eluent compositions on a small set of model compounds.

EXPERIMENTAL

Chemicals, equipment and procedures were as described previously [2]. THF was HPLC grade from FSA Scientific Apparatus (Loughborough, UK).

The retention indices were calculated by linear interpolation of the retentions of the model compounds between the log k' of the alkyl aryl ketones, which were assigned retention index values of carbon number \times 100. Retention index $I = 100 n +$ 100(log k'_x - log k'_n)/(log k'_{n+1} - log k'_n) in which k'_x = capacity factor of model compound and k'_n and k'_{n+1} = the capacity factors of the alkyl aryl ketones, containing n and $n + 1$ carbon atoms, respectively, which were eluted immediately before and after the analyte. For compounds which eluted more rapidly than acetophenone the scale was extrapolated from the results for acetophenone and propiophenone.

RESULTS AND DISCUSSION

The retention times and capacity factors were determined for a series of alkyl aryl ketones from acetophenone to heptanophenone and for three sets of model compounds, alkylbenzenes from benzene to butylbenzene (five compounds), substituted benzenes (nineteen compounds) and terminal substituted propylbenzenes (four compounds). The last set was used as model for aliphatic substitution as previously it has been demonstrated that there were no significant interactions between

TABLE I

CAPACITY FACTORS OF MODEL COMPOUNDS ON ELUTION WITH THF-BUFFER ELUENTS

functional groups in this position and the phenyl ring [8]. A range of THF-pH 7.0 phosphate buffers eluents (20:80 to 60:40) gave capacity factors in the range 0.2-100 (Table I). As with the previous studies [2], it was impractical to measure the retentions of carboxylic acids or of aliphatic amines as these compounds were often partially or completely ionised in the pH 7.0 buffer. The stationary phase was Spherisorb ODS-2 from the same batch as the earlier work [2] to ensure a uniform set of retention data.

Fig. 1. Relationship between log k' and carbon number \times 100 for the alkyl aryl ketones at different % of THF.

Unlike the previous studies [2] using methanol or acetonitrile as the organic modifier, the relationships between the logarithms of the capacity factors of the alkyl aryl ketones and their carbon numbers was not linear as acetophenone was eluted more rapidly than predicted in each eluent (Fig. 1). A similar but less pronounced effect was also observed in the preliminary studies [7] on a Spherisorb ODS column. This effect might be caused by a difference between the empirical value for the column void volume, measured using aqueous sodium nitrate solution, and the effective void volume experienced by the alkyl aryl ketones. A linear interpolation between the logarithms of the capacity factors of the alkyl aryl ketones was therefore used to calculate the retention indices of the model compounds (Table II). The retention indices were similar to those obtained previously [7] on a Spherisorb ODS column (e.g., at 30% THF, previous values in parenthesis: methyl benzoate, $I = 887$ (896); nitrobenzene, $I = 915(931)$; toluene, $I = 1061(1049)$. However, the differences in the index values were greater than the expected experimental error of ± 10 units [2] indicating a small but significant difference in the retention properties of the two stationary phases.

In order to determine the coefficients of the parent index value for benzene (I_P) , the experimental retention indices were related to the percentage of THF (x) using a quadratic least-squares correlation to give the equation $I_P = 0.0243x^2 + 1.037x +$ 913 over the composition range 20–60% THF. As in the earlier studies, extrapolation outside the empirically determined range is likely to be inaccurate. Using this equation the predicted retention indices of benzene in the different eluent compositions were calculated (Table II). They showed a close correspondence to the empirical values and will be used as the reference values to calculate the effects of the substituents on retention. The values of the parent indices in the THF-buffer eluents (I_p) $= 944 - 1063$) are generally higher than those found earlier for the parent indices of benzene in methanol-buffer ($I_P = 888-999$) or acetonitrile-buffer eluents ($I_P = 910-$ 963) [2].

The rentention indices for the alkylbenzenes (benzene, toluene, ethylbenzene and n -propylbenzene) increased with the chain length as expected (Table II) but the

TABLE II

RETENTION INDICES OF MODEL COMPOUNDS ON ELUTION WITH THF-BUFFER ELUENTS

^a Calculated as $I_p + I_{i, PhCH_2R} + 3 \times 100$.

 b Defined as carbon number \times 100.

increases for each additional methylene unit were often smaller than the theoretical value of 100 units expected for the increments between homologues (Table III). The mean value for the addition of a methylene group to the benzylic carbon was significantly smaller (91 units) than the increment for the other methylene groups (95-98 units). A similar effect was observed earlier in studies with methanol- and acetonitrilecontaining eluents, when the increments were in the range 92-112 units, except for the addition of a methylene group to a benzylic carbon, 87-95 units [9]. These results suggest that the methylene increment for the alkylbenzenes in a THF-buffer eluent

TABLE III

RETENTION INDEX INCREMENTS FOR THE ADDITION OF A METHYLENE GROUP TO AN ALKYLBENZENE

differs from that of the alkyl aryl ketones. From the Martin equation, the methylene increment should be constant irrespective of the functional groups. However, similar differences have been observed previously, particularly between aliphatic and aromatic homologous series, as in a recent study comparing homologous nitroalkanes with 2-alkanones and alkyl aryl ketones [10]. In the present study, an interaction index term was therefore included for the addition of a methylene group to a benzylic carbon ($I_{I,PhCH_2R}$ = -14) similar to the correction in methanol and acetonitrile eluents $(I_{\text{I,PhCH}_2R} = -12)$ [9]. This interaction value ignores the low value for alkyl substitution on benzene but would compensate by slightly over-correcting for substitution of toluene. Using this interaction value the predicted retention indices can be calculated for *n*-propylbenzene (Table II) and these will be used as the reference point for aliphatic substitution.

The differences in the effect of methylene groups on retention could pose problems for retention prediction systems and imply that additional correction factors might be required for precise results on different homologous series. However, the discrepancies are small compared to the effects of many of the substituents and are similar to the uncertainty in the empirical measurement of retention indices. Thus, they can probably be safely ignored in most predictions.

Substituent index increments

Using the calculated parent index values for benzene (I_P) and *n*-propylbenzene $(I_{R-H} = I_P + I_{I,PhCH_2R} + 3 \times 100)$ the changes in the retention caused by the presence of each aromatic and aliphatic substituent were calculated as the retention index increments between the substituted model compound and the corresponding parent hydrocarbon $(\delta I = I_{Ar-X} - I_{Ar-H}$ or $\delta I = I_{R-X} - I_{R-H}$) (Table IV). In each case the increments changed systematically with eluent composition. As with the other eluent modifiers, there were large differences between the same group as an aliphatic or aromatic substituent *(e.g, in 40% THF,* δI *for Ar-Br = 92 and R-Br = -29).*

Although, as expected, the difference between the amino group $(-NH₂)$ and the N-ethyl group (-NHEt) in THF-buffer (20:80) was about 200 units, the difference

RETENTION PREDICTION IN RP-HPLC. VII.

TABLE IV

Substituent	Hansch constant. π [12]	Retention index increment THF (%)				
		$Ar-X$ ($\delta I = I_{Ar.X} - I_{Ar-H}$)				
SO_2NH_2	-1.82	-281	-301	-335	-394	-469
CONH,	-1.49	-335	-381	-434	-481	-596
CONHCH ₃	-1.27	-323	-370	-416	-494	-570
$CON(CH_2)$,	\overline{a}	-339	-385	-430	-494	-570
NH,	-1.23	-208	-213	-230	-264	-301
NHCOCH,	-0.97	$\overbrace{}$	-278	-333	-399	-486
OH	-0.67	-123	-141	-177	-231	-296
CHO	-0.65	-149	-163	-188	-218	-251
OCOCH ₃	-0.64	—	-125	-149	-181	-250
CN	-0.57	-115	-131	-160	-199	-242
COCH, ⁴	-0.55	-144	-166	-194	-222	-263
NO ₂	-0.23	-37	-51	-88	-133	-179
OCH ₃	-0.02	-15	-27	-49	-74	-94
CO,CH ₃	-0.01	-49	-79	-113	-151	-190
н	0.00	$\bf{0}$	$\mathbf{0}$	0	$\mathbf{0}$	$\bf{0}$
NHCH, CH,	0.08	-1	10	$\bf{0}$	-26	-59
F	0.14	28	26	7	-12	-24
CH ₃	0.56	95	95	89	92	102
Cl	0.72	118	100	76	52	51
Вr	0.86	142	120	92	67	54
Ph	1.96		240	201	161	132
(Naphthalene)			146	107	76	51
$R-X(\delta I = I_{R-X} - I_{R-H})$						
OН	-1.64		-435	-491	-557	-623
CN	-1.27		-318	-352	-407	-462
C1	0.06		-25	-59	-85	-116
Br	0.20		$\overline{2}$	-29	-56	-84

RETENTION INDEX INCREMENTS FOR SUBSTITUENTS ON AN AROMATIC RING IN THF-**BUFFER ELUENTS**

increased to 250 units at higher proportions of modifier, suggesting that the influence of the eluent on primary and secondary amino groups differed and that different substituent index coefficients are necessary. A comparison of the increments for the carboxamide $(-\text{CONH}_2)$, N-methylcarboxamide $(-\text{CONHMe})$ and N,N-dimethylcarboxamide $[CON(Me)_2]$ groups showed an even more marked effect. At each eluent composition the increments for all three functional groups were very similar $(i.e.,$ THF-buffer (20:80), $I_{S,Ar-X} = -335$, -323 and -339 units, respectively, and at THF-buffer (60:40), $I_{s,Ar-X} = -596$, -570 and -570 units) despite the presence of the additional methyl and dimethyl groups, which might have been expected to make the subsituted carboxamide groups more hydrophilic. The retention prediction system will therefore have to regard primary, secondary and tertiary amides as different groups (with different substituent index coefficients), rather than simply considering them as primary amides, which have been altered by alkyl substitution. During the evaluation of the retention index expert system (CRIPES) this simpler approach was used for the prediction of the indices of N,N-dimethylbenzamide and N-methylbenzamide in methanol and acetonitrile eluents [3]. However, there were large differences between the predicted and experimental values and it was clear that the substituted amides would also need to be distinguished in those eluents.

Using the retention index increments for each substituent the coefficients of the quadratic relationship between the increment and the percentage of THF in the eluent

TABLE V

COEFFICIENTS OF SUBSTITUENT INDEX EQUATIONS FOR SUBSTITUENTS

 $I_{A_{r-1}}$ or $I_{R-x} = ax^2 + bx + c$, $x = \frac{9}{6}$ THF in eluent.

^a Values of core functional groups excluding alkyl groups (methyl = 100, ethyl = 200).

Fig. 2. Relationship between retention increments or calculated substituent indices for selected aromatic substituents and the proportion of THF in the eluent. Points indicate experimental values (retention index increment, δf) and the lines are calculated substituent indices $(I_{S, A\tau} | X)$ based on the coefficients in Table V.

were calculated (Table V). For substituents containing aliphatic, methyl or ethyl groups (such as $-COMe$, $-CONMe₂$ and $-NHEt$) the corresponding coefficients for the core functional group $(-CO₋, -CON =$ and $-NH₋$) were determined by subtracting 100 or 200 units as appropriate. These coefficients can then be used as general terms for the prediction of the retention indices of substituted functional groups (such as $-NH$ butyl) by adding the appropriate alkyl chain substituent contribution $(I_{S,R})$. From the coefficients in Table V the substituent indices $(I_{S, A r-X}$ and $I_{S, R-X}$) can be calculated and these corresponded closely with the empirical retention increments (for examples see Fig. 2).

The substituents indices in the three different modifiers can now be compared to identify those groups which are particularly susceptible to differences in the properties of the eluents (Table VI). As in a typical optimisation procedure approximately isoeluotropic conditions were selected for comparison. These were chosen to give similar capacity factors for acetophenone (methanol-buffer (50:50), $k' = 3.23$; acetonitrile-buffer (40:60), $k' = 2.91$; THF-buffer (30:70), $k' = 3.48$). The parent index values for benzene differed in each eluent ($I_P = 913, 927$ and 966, respectively). In contrast, the value for the phenyl substituents decreased $(I_{S,Ar-Ph} = 305, 271$ and 241, respectively). For many of the substituents the differences between the eluents were less than 30 units and the aromatic nitro $(I_{\rm S, A}I_{\rm CO_2} = -54, -54, -56$ and -56) and chloro groups $(I_{S,Ar-Cl} = 105, 98$ and 92) had almost identical indices. However, these comparisons are not general and would alter with the strength of the eluents. For example, the substituent index of the chloro group (Fig. 3) changed markedly with THF composition, whereas in methanol and acetonitrile eluents it was almost constant with eluent composition [2].

The largest differences between modifiers was found for the phenolic hydroxy group $(I_{S,AT-OH}$, methanol-buffer = -229; acetonitrile-buffer = -243 and THF-

TABLE VI

COMPARISON OF SUBSTITUENT INDICES IN METHANOL-, ACETONITRILE- AND THF-BUFFER-CONTAINING ELUENTS

 $I_{\text{S.Ar-X}}$ and $I_{\text{S.R-X}}$ based on coefficients in [2,8] and Table V.

Eluents selected to give similar capacity factors for acetophenone; 50:50 methanol-buffer, $k' = 3.23$; 40:60 acetonitrile-buffer, $k' = 2.91$; 30:70 THF-buffer, $k' = 3.48$

buffer $= -142$). In contrast, the aliphatic hydroxyl group behaved very differently $(I_{\text{S.R}-\text{OH}} = -362, -459 \text{ and } -434$, respectively). Although the values for the carbomethoxyl substituent (CO_2CH_3) were relatively similar in the three isoeluotropic eluents the values became markedly different in eluents with higher proportions of modifier (Fig. 3). These changes emphasise that differences in the elution order between modifiers may be significantly dependent on the composition of each eluent and it may not be sufficient to rely on comparisons of a single set of isoeluotropic eluents to achieve the optimisation of a separation.

However, even relatively small differences of less than 50 units in different eluents can have a marked influence on the relative retentions of the components of a mixture. A typical pair of compounds, *e.g.,* methyl benzoate and nitrobenzene, have predicted retention indices of 903 and 859 units, respectively, in a methanol-buffer (50:50) eluent but their elution would be reversed to $I = 886$ and 910 units in THFbuffer (30:70). This predicted change would mirror changes reported previously for these compounds by Tanaka *et al.* [11] on an ODS-bonded Hypersil column.

Many workers have used HPLC retentions for the approximate determination of octanol-water distribution constants ($log P$) and close correlations have often been

Fig. 3. Comparison of the substituent indices of the aryl chloro group (Ar-CI, solid symbols) and the carbomethoxy group (Ar-CO₂CH₃, open symbols) in THF-buffer eluents (\circ and \bullet), methanol-buffer eluents (\Box and \Box) and acetonitrile eluents (\triangle and \blacktriangle) (values from ref. 2).

found, particularly amongst members of a homologous or pseudohomologous group of compounds (such as the alkylsubstituted barbiturates). The substituent indices were therefore compared with the corresponding Hansch substituent constants (π values [12], see Table IV). Across the range of aromatic substituents there was a close correlation (Fig. 4). Unlike the corresponding comparison in methanol and acetonitrile eluents [2], the value for the aromatic hydroxyl group in THF-buffer correlated with its π value. However, the carboxamide (as with the other eluent modifiers) and sulphonamide groups were outliers, although in different directions.

Fig. 4. Relationship between the magnitude of the calculated substituent indices of aromatic substituents in THF-buffer (30:70) eluent and Hansch π values (from Table IV).

CONCLUSIONS

The parent retention indices for benzene and substituent indices for 21 aromatic substituents and 4 aliphatic substituents have been determined and expressed as the coefficients of quadratic equations covering an eluent composition range from 20 to 60% THF. The magnitude of the substituent indices corresponded closely to the Hansch π values for the groups.

ACKNOWLEDGEMENTS

The authors thank Liaoning University, China, for financial support and study leave to R.W. and Phase Separations Ltd. for a gift of Spherisorb ODS-2.

REFERENCES

- 1 G. L. Glach and L. R. Snyder (Editors) *Computer-Assisted Method Development for High-Performance Liquid Chromatography,* Elsevier, Amsterdam, 1990; J. *Chromatogr.,* 485 (1989).
- 2 R. M. Smith and C. M. Burr, *J. Chromatogr.,* 475 (1989) 57.
- 3 R. M. Smith and C. M. Burr, *J. Chromatogr.,* 485 (1989) 325.
- 4 R. M. Smith and C. M. Burr, J. *Chromatogr.,* 550 (1991) 335.
- 5 J. C. Berridge, *Techniques for the Automated Optimisation of HPLC Separations,* Wiley, Chichester, 1985.
- 6 A. Ahuja, *Selectivity and DetectabilityOptimisations in HPLC,* Wiley, New York, 1989.
- 7 R. M. Smith, G. A. Murilla and C. M. Burr, *J. Chromatogr.,* 388 (1987) 37.
- 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 N. Finn, J. *Chromatogr.,* 518 (1991) 75.
- 11 N. Tanaka, H. Goodell and B. L. Karger, J. *Chromatogr.,* 158 (1978) 223.
- 12 C. Hansch and A. Leo, *Substituents Constants for Correlation Analysis in Chemistry and Biology,* Wiley, New York, 1979.