This item was submitted to Loughborough's Research Repository by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Retention prediction in RP-HPLC

PLEASE CITE THE PUBLISHED VERSION

PUBLISHER

© Christina M. Burr

PUBLISHER STATEMENT

This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 2.5 Generic (CC BY-NC-ND 2.5) licence. Full details of this licence are available at: http://creativecommons.org/licenses/by-nc-nd/2.5/

LICENCE

CC BY-NC-ND 2.5

REPOSITORY RECORD

Burr, Christina M.. 2019. “Retention Prediction in RP-HPLC”. figshare. https://hdl.handle.net/2134/27346.

Loughborough University

This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (https://dspace.lboro.ac.uk/) under the following Creative Commons Licence conditions.

creative commons

COMMONS DEED

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

Attribution. You must attribute the work in the manner specified by the author or licensor.

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the Legal Code (the full license).

Disclaimer

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/

BLDSC no:- DX 88177

LOUGHBOROUGH
UNIVERSITY OF TECHNOLOGY
LIBRARY

AUTHOR/FILING TITLE
BURR, C M

ACCESSION/COPY NO.
040013084

VOL. NO. | CLASS MARK
LOAN COPY

- 5 JUL 1991
25 APR 1997
3 JUL 1992
14 APR 2000
- 2 JUL 1993
- 6 OCT 2000
- 4 OCT 1996
18 MAR 1997

040013084 X

THIS BOOK WAS BOUND BY
BADMINTON PRESS
18 THE HALCROFT
SYSTON
LEICESTER LE7 2LD
0533 602917

RETENTION PREDICTION IN RP-HPLC

by

CHRISTINA MARY BURR

A Doctoral Thesis

submitted in partial fulfilment

of the requirements for the award of

DOCTOR OF PHILOSOPHY

of

Loughborough University of Technology

October 1988

Supervisor Dr. R.M. Smith,

B.Sc., M.Sc.,Ph.D., C.Chem. F.R.S.C

Department of Chemistry

Loughborough University of Technology

© by Christina M. Burr, 1988

Loughborough University
of Technology Library
Date Oct. 89
Class
Acc. No. 040013084

ACKNOWLEDGMENTS

I would like to thank my supervisor Dr. R.M. Smith for all his advice and encouragement throughout the period of the study.

I would also like to thank the technical staff for their assistance in particular John Kershaw and John Swithenbank.

I am also grateful for the help and companionship of my friends and colleagues in particular Rajani Shrestha and Paul Scullion. And special thanks to Dave.

Finally I would like to thank SERC for my grant, Phase Separations, U.K. for the gift of the stationary phase and Dr. L.R. Snyder of LCResources for the Drylab I program.

ABSTRACT

A method of calculating the RP-HPLC retention indices, based on the alkylarylketone scale, has been developed. The retention indices are calculated from the molecular structure of a compound as the sum of the parent contribution, the parent index, the substituent contributions, the substituent indices, and terms to account for the interactions between substituents, the interaction indices.

The substituent contributions have been determined for 12 aliphatic functional groups and 14 aromatic functional groups over a range of methanol and acetonitrile eluents. Interactions have been studied in 73 disubstituted aromatic compounds to obtain empirical interaction indices.

The changes, with the proportion of modifier in the mobile phase, of the parent contribution, the substituent contributions and the interaction terms have been fitted to quadratic equations. The corresponding index values used for the prediction of retention indices are obtained from these equations. The calculated retention indices can therefore be described by quadratic regression equations. Using the relationship between capacity factor and carbon number for the alkylarylketone standards it is also possible to estimate a capacity factor from the retention index of a compound. This enables the expected resolution between two compounds to be determined.

An expert system program, CRIPES, Chromatographic Retention Index Prediction Expert System, has been written to implement the prediction system. This was extended to calculate the retention between two compounds. The accuracy

of the prediction system was examined by comparing the experimental and calculated retention indices of a number of compounds.

A commercial retention prediction/ optimisation program (Drylab-I) was also briefly examined.

CONTENTS

		Title Page	
		Acknowledgements	
		Abstract	
		Contents	
CHAPTER 1		Introduction and literature survey	1
	1.1	Introduction	1
	1.2	Retention prediction in RP-HPLC	2
	1.2.1	Prediction based on size and shape parameters	3
	a	Carbon number	3
	b	Topological indices	5
	c	Size and shape parameters of polynuclear aromatic hydrocarbons (PAH)	7
	d	Van der Waals volume or radius	8
	e	Other size and shape parameters	9
	1.2.2	Prediction based on octanol-water partition coefficients	9
	1.2.3	Prediction based on aqueous solubility	13
	1.2.4	Prediction based on solvatochromic parameters	14
	1.2.5	Prediction methods using group contributions	16
	a	Group contributions using capacity factors	17
	b	Group contributions using retention indices	22
	c	Prediction based on interaction indices	23

	1.2.6	Miscellaneous retention prediction methods	24
	a	Mobile phase relationships	24
	b	Prediction based on thermodynamic parameters	25
	c	Prediction using the electronic interaction index	25
	d	Computer based prediction methods	26
	1.3	Use of retention indices in RP-HPLC	26
	1.3.1	Retention indices in retention prediction	29
	1.4	Introduction to current work	30
CHAPTER 2		Experimental	33
	2.1	Chemicals	33
	a	Mobile phase components	33
	b	Retention index standards	33
	c	Model and test compounds	33
	d	Void volume marker	33
	2.2	HPLC equipment	34
	2.3	Computing	34
	2.4	Experimental procedure	35
	2.4.1	Mobile phase preparation	35
	2.4.2	Solute preparation	35
	2.4.3	Void volume solution	36
	2.4.4	Column testing	36
	2.4.5	Experimental method	37
	2.5	Calculations	37
	2.5.1	Capacity factors	37
	2.5.2	Retention indices	38
	2.5.3	Parent indices	38
	2.5.4	Substituent indices	38

	2.5.5	Interaction increments	39
	2.5.6	Retention prediction	39
CHAPTER 3		Reproducibility and robustness of capacity factors and retention indices	40
	3.1	Method protocols	42
	a	Temperature	42
	b	Buffer pH	43
	c	Stationary phase	44
	d	Experimental procedure	44
	e	Calculations	45
	3.2	The determination of column void volume and the effect on capacity factors and retention index	46
	3.2.1	Investigation of the dependence of the retention of sodium nitrate on its concentration	47
	3.2.2	Investigation of the dependence of the retention time of sodium nitrate on the eluent composition	48
	3.2.3	Calculation methods for the determination of column void volume	51
	a	Calculation of column void volume using the Berendson method	52
	b	Calculation of column void volume using an iterative method	53
	3.2.4	Reproducibility of measured retention time of sodium nitrate	56
	3.2.5	Effect of changes in the column void volume value on capacity factors and retention indices	57

3.3 Dependence of capacity factors and retention indices on the buffer concentration of the aqueous component of the mobile phase 64
3.4 Reproducibility of capacity factors 68
3.4.1 Short term reproducibility of capacity factors 69
3.4.2 Long term reproducibility of capacity factors 71
a Capacity factors on each column 72
b Capacity factors of benzene, phenol and toluene 72
3.5 Calculation of retention indices 76
3.5.1 Linearity of log k'- carbon number relationships for the alkylarylketones 77
3.5.2 Relationship between slope and intercept of correlation curves and eluent composition 83
3.6 Reproducibility of retention indices 86
3.6.1 Short term reproducibility 86
3.6.2 Long term reproducibility 88
CHAPTER 4 Parent index and substituent indices for single substituents on aliphatic side chains 91
4.1 Definition and calculation of parent index values 92
4.2 Retention behaviour of aliphatic substituted benzenes 93
4.2.1 Capacity factors of model compounds 96
4.2.2 Retention indices of model compounds 100

4.2.3 Retention behaviour of benzylamine, 2-phenylethylamine and 3-phenyl-1-propylamine 106
4.3 Substituent indices for aliphatic substituents 111
4.3.1 Methylene group increment for alkyl side chains 111
4.3.2 Substituent indices of substituents on the aliphatic side chain 114
4.4 Effect of unsaturation in the alkyl chain on retention index 121
4.5 Relationship between *SI* and structural parameters 123
4.5.1 Relationship between *SI* and π. 123
4.5.2 Relationship between *SI* and aqueous solubility 126
4.6 Interactions of substituents with the aromatic ring 132
4.7 Use of substituent index equations and interaction index equations 138
CHAPTER 5 Investigation of the retention aliphatic compounds with isomeric substituent and disubstituted aliphatic compounds 139
5.1 Examination of the effect of chain branching on the retention of alkylbenzenes 139
5.1.1 Substituent contributions of the methyl/methylene groups in isomeric alkylbenzenes 141
5.1.2 Relationship between retention

increment and structural parameters 144

5.2 Examination of the retention of isomeric phenylpropanols 145

5.2.1 Retention increments calculated for hydroxyl groups 147

5.3 Retention behaviour of selected compounds with two substituents on the alkyl chain 151

CHAPTER 6 Substituent indices of aromatic substituents 155

6.1 Retention indices of mono-substituted model compounds 155

6.2 Substituent indices for single aromatic substituents 161

6.3 Relationship between substituent index and physico-chemical properties 165

6.3.1 Relationship between *SI* and octanol-water partition coefficients 168

CHAPTER 7 Retention behaviour of disubstituted benzenes 172

7.1 Retention indices of disubstituted benzenes 175

7.2 Interaction increments for pairs of aromatic substituents 186

7.2.1 Interaction indices 195

7.2.2 Interactions in disubstituted compounds 195

a Interactions with hydroxyl and carbonyl containing compounds 200

b Interactions in hydroxyl-hydroxyl and

hydroxyl-amino compounds 200

c Interactions in hydroxyl-nitro compounds 201

d Interactions in compounds without hydrogen bonding groups 202

7.3 Relationship between interaction indices and studies on octanol-water partition coefficients 204

7.3.1 Possible application of the above approach to the current study 209

a Relationship between interaction increments and Hammett constants for hydroxyl substituted compounds 209

b Relationship between Leo method of correcting for interaction and interaction increments 211

7.4 Application of interaction indices 214

CHAPTER 8 Implementation of the prediction method using an expert system 216

8.1 Use of computer prediction methods in HPLC 216

8.2 Introduction to expert or knowledge based information systems 219

8.3 Use of expert systems in chromatography 221

8.4 Introduction to VP-expert 225

8.5 Implementation and development of CRIPES 229

8.6 Calculation of resolution 237

8.7 Testing of CRIPES 241

a Phenacyl halides 247

	b	Compounds with amino groups	247
	c	Amide test compounds	248
	d	Secondary amine	249
	e	Aliphatic substituted compounds	249
	f	Other compounds	249
8.8		Comparison of a commercial retention prediction system, Drylab-I and CRIPES	250
CHAPTER 9		Extending the prediction system using published data	257
9.1		Comparison of retention indices determined previously with values calculated in this study	259
9.2		Retention indices calculated from published data	261
9.3		Substituent increments using retention indices calculated from published data	262
9.4		Interaction increments using retention indices calculated from published data	266
9.5		Use of octanol-water partition data	268
9.6		Use of group contribution data	270
9.7		Expanding CRIPES	270
CHAPTER 10		Conclusion	271
		Future work	273
		REFERENCES	274
Appendix 1		Preliminary studies on the use of retention indices in tetrahydrofuran (THF)	287
Appendix 2		Published work	295

CHAPTER 1

INTRODUCTION AND LITERATURE SURVEY

1.1 INTRODUCTION

The aim of this study was to develop a method to calculate the reversed-phase high performance liquid chromatography (RP-HPLC) retention index of a compound from its molecular structure. If the retention of two (or more) compounds could be calculated, the separation of the compounds could then be optimised.

Most current optimisation procedures require the retention of a compound to be measured at a minimum of two eluent compositions, or in two organic modifiers, and frequently at considerably more than two eluent compositions. From this information the retention at another eluent composition can be obtained by interpolation between the eluent compositions or extrapolation beyond the values. It has also been possible to "predict" the retention of a compound from a gradient elution run although this is also a form of extrapolation or interpolation. These essentially empirical methods do not require a knowledge of the molecular structure of the compounds being separated.

However, the structures of the compounds to be chromatographed are frequently known. Relationships between the retention in RP-HPLC and the structure of compounds have been noted for many different structural descriptors. These quantitative structure-retention relationships (QSRR) have recently been reviewed by Kaliszan[1]. A literature survey of the prediction methods using structural parameters is

presented here.

One of the problems with RP-HPLC retention prediction is the dependence of the absolute retention on the column used. Although the most frequently used measure of retention, the capacity factor, can compensate for differences in flowrate and column dimensions there can be considerable differences between nominally equivalent stationary phases and even different batches of the same stationary phase. Various workers have therefore suggested the use of interpolated retention index scales in which the retention is expressed relative to a set of standards. These scales have been shown to be robust to changes in the HPLC system. Relatively few prediction methods have been developed using retention indices but their use should result in more generally applicable prediction methods. A survey of the use of retention indices in RP-HPLC is also given in this chapter.

1.2 RETENTION PREDICTION IN RP-HPLC

There have been many different approaches to retention prediction in RP-HPLC the simplest of which are those in which the retention is calculated by extrapolation or interpolation between experimental data, these methods will not be discussed in this section. More useful prediction methods have been based on additive structural properties of the solute. These prediction methods generally excluded any specific interactions of a compound with the hydrocarbonaceous stationary phase, assuming that the mobile phase has a dominant role in the retention process. This

solvophobic theory of retention was proposed by Horvath[2,3] and has been widely used. Other retention mechanisms have also been proposed but few have been used in retention prediction because of the more complex parameters which would be required.

Although all the methods discussed below have been suggested for retention prediction, few of the methods have actually been used to predict the retention of "unknown" compounds. In each case there was an empirical, usually linear, relationship between the structural parameter and retention in reversed-phase high performance liquid chromatography. In most cases the system was calibrated using a set of standards to obtain a regression equation and the retention of an "unknown" compound was calculated from the structural feature either by extrapolation or interpolation. Frequently multi-parameter equations have been used based on combinations of different structural properties.

1.2.1 Prediction Based on Size and Shape Parameters

Many prediction methods have been based on the relationship between retention and size or shape descriptors. The different size and shape descriptors which have been used will be discussed in this section.

a) Carbon Number

The simplest descriptor of size is the carbon number. A linear relationship between carbon number and log capacity factor (k') for the members of a homologous series has been

described for many different homologous series[4,5]. This relationship has also formed the basis of most retention index scales in both LC and GC (see Section 1.3). Prediction based purely on this parameter is limited to the members of the same homologous series[6-9]. To predict the retention of a compound it is necessary to measure the retention of two or more members of the same homologous series and calculate the retention either by extrapolation or interpolation. Dufek[7-9] described a method to calculate the retention of any member of a homologous series at any temperature and any eluent composition from the retention of two other members of the same homologous series. Viseras *et al.* extended the prediction method proposed by Dufek to gradient elution chromatography[10].

The carbon number alone was found to be insufficient to account for the retention of 54 alkylbenzenes and polymethylbenzenes[11], the number of hydrogen atoms was also not successful in describing the retention. Schabron[12] combined the use of the carbon number with the number of double bonds to calculate the capacity factors of a number of alkylphenols. Schronk *et al.*[13] also combined carbon number with the number of double bonds and rings to describe the retention of nitrogen heterocyclic compounds. A number of multi-parameter regression equations were derived but they were not used for prediction.

Golovnya *et al.*[14,15] proposed a "universal" equation (1.1) for calculating the retention (Z) based on the number of a compound in a homologous series (m) rather than the carbon number.

$$Z = \alpha + \beta_m + (\gamma \log m)m^{-1} + \varepsilon[(m-2)^2 + 0.1]^{-1} \qquad 1.1$$

Each of the coefficients was determined experimentally for a homologous series. The equation was used for the calculation of the retention of 2,4-dinitrophenylhydrazone derivatives of carbonyl compounds. As with the carbon number this equation could only be used for members of a homologous series.

b) Topological Indices

A more useful and more general size and shape descriptor which has been suggested for retention prediction is the molecular connectivity index (X). Molecular connectivity indices are based on graph theory and provide a description of the shape of the molecule. Various degrees of molecular connectivity indices have been described by Kier and Hall[16], who also very briefly discussed on the possibility of their use in chromatography. The most commonly used X, the first order molecular connectivity index, ${}^1X$, is based on the relationship of the nearest member in the molecule and is calculated for each bond according to the equation[17]:-

$$ {}^1X = \Sigma\ (\delta_i \delta_j)_m^{-0.5} \qquad 1.2 $$

where δ_i and δ_j were the number of atoms to which the atoms at either end of the bond, s, were attached. Attachment to hydrogen atoms is ignored. However, this first order molecular connectivity index would not account for the different properties of heteroatoms because it counts only the number of atoms attached, not the type of atoms to which they are attached. Linear relationship between molecular

connectivity index and log k' has been reported by several workers[18-22]. However, few workers have extended the approach to predicting the retention of compounds from their calculated X values[19-21]. Molecular connectivity indices have been found to be insufficient to account for the elution order of branched chain alkanes[23], alkylbenzenes[11,24], and diols[25]. The problem of predicting the elution orders of isomeric compounds has led to the use of combinations of molecular connectivity indices with other topological indices and structural parameters. Burke et al.[26] used combinations of 11 different topological indices to predict the retention of a range of liquid coal components. The relative contribution of each index was determined by a multi-parameter regression equations and the results were reasonably successful.

Funasaki et al.[27] combined the use of molecular connectivity indices with the octanol-water partition coefficients. Molecular connectivity indices have been used by Jinno and co-workers[28-30] in combination with octanol-water partition coefficients (log P) and Van der Waals radii or volumes. The approach by these workers will be discussed later.

An alternative topological index, the Weiner Index, has also been used to predict retention. The Weiner Index is a descriptive parameter based on molecular graph theory which was used by Adler for the calculation of HPLC retention of polynuclear aromatic hydrocarbons[31]. Adler observed a non-linear relationship between log capacity factor and the Weiner Index which contrasted with the linear relationships reported for other topological indices. Noel reported a higher correlation between the log capacity factor and the

Weiner Index than between the capacity factor and the molecular connectivity index[25]. A combination of the Weiner index and octanol-water partition coefficients was suggested to predict the capacity factors of diols[25].

c) Size and Shape Parameters of Polynuclear Aromatic Hydrocarbons (PAH)

The retentions of PAH have been shown to be predominantly determined by the size and shape of the molecule[32]. Various workers have suggested the use of a length/breadth ratio to predict the retention of these compounds. The definition and method of calculating the length/breadth ratio has differed among the different groups but success has been reported by several workers, Jinno and Kawasaki[33,34], Rohrbaugh and Jurs[35], and Wise[36]. Jinno and Kawasaki[33] combined the length/breadth parameter with the Chromatographic Correlation Parameter (F) proposed by Schabron[37]. This chromatographic correlation parameter was calculated using the equation[12,34,37].

F = (no. of double bonds) + (no. of 1° and 2° carbon atoms) - 0.5 for each non-aromatic ring

1.3

Jinno and Kawasaki[33] reported that the correlation between retention and F was better than the correlation between retention and molecular connectivity indices. Hasan and Jurs[38] have also described a method of calculating the retention indices of PAH based on a multiple parameter equation containing a number of terms describing the

molecular structure.

d) Van der Waals Volume or Radius

The Van der Waals radius or volume has been used by several groups for retention prediction, although usually as one of several parameters rather than as a single descriptor. Linear relationships between log capacity factor and Van der Waals volume were reported by Hanai *et al.*. Although the slopes were different for different groups of compounds. They attempted to predict the retention of a variety of compounds by combining the Van der Waals volume with delocalisation energy effects[39-41].

Van der Waals volumes were also used by Jinno and co-workers in combination with other structural parameters (X and octanol-water partition coefficients) to calculate the capacity factors of isomeric alkylbenzenes[28,29,33]. The Van der Waals volume was also used as the basis of a prediction method by Yugi *et al.*[42]. Firstly a contribution of solute hydrophobicity (log k'_o) to retention was calculated from the Van der Waals volume using the equation.

$$\log k'_o = \alpha V_\omega + \gamma \qquad 1.4$$

A hydrophilic parameter (D) for a solute was then calculated as the difference between the measured capacity and the calculated log k'_o. A linear relationship between D and the mobile phase composition was used to calculate an extrapolated value (D_ω) at 0% organic modifier. Overall an equation was proposed to calculate the retention of a solute at a specified eluent composition.

$$\log k' = \alpha V_w - \Sigma D_{i,j} + \gamma \qquad 1.5$$

where $D_{i,j}$ were the hydrophilic group contributions calculated as above. The method was successful in calculating the retention of 4 substituted benzenes in which the substituents did not interact but was not successful in calculating the retentions when substituents interacted.

e) Other Size and Shape Parameters

A few other size and shape parameters have been used to predict the retention of compounds. The General Index of Molecular Complexity (GIMC) was suggested by D'Amboise[43]. The GIMC was another parameter which attempts to account for branching, size, the degree of branching and the presence of heteroatoms but the method was only found to be applicable to members of a single homologous series.

Mockel *et al.*[44-47] related the retention of various classes of compounds to the total surface area of the solute. It was suggested that this could provide a method of calculating the retention of other compounds although it was not used for this purpose but to rationalise the different retention characteristics of classes of compounds.

1.2.2 Prediction Based on Octanol-Water Partition Coefficients

The octanol-water partition coefficient (log *P*) of a compound is the distribution coefficient of that compounds between an octanol phase and an aqueous phase. It has

frequently described as the "lipophilicity" or "hydrophobicity" of a molecule and is dependent on the molecular structure. Traditionally octanol-water partition coefficients have been obtained using the "shake-flask" method in which a known amount of the compound is shaken in an octanol-water mixture for a fixed period of time and the concentration determined in one or both of the phases. Hansch[48] proposed a method of calculating the octanol-water partition coefficient of a compound as the sum of the parent log *P* plus the substituent contributions (π).

$$\log P = \log P_{PhH-H} + \Sigma\pi \qquad 1.6$$

A large number of π values have been tabulated[48]. In this form the calculation does not account for any interactions occurring between substituents in a multiply substituted molecule. More recently methods to account for the interactions have been described[49-52]. An alternative calculation method has been proposed by Rekker[53] in which the log *P* was calculated as the sum of fragmental constants, f, which were derived in a different manner to the Hansch constants. The Hansch method has probably been more widely used.

The similarity between the distribution processes in an octanol-water system and the RP-HPLC column led to suggestions that the HPLC retention could be related to the value of the octanol-water partition coefficient[1,54]. Linear relationships between log *k'* and log *P* have been reported for many groups of compounds[1,54,55]. However, it has been noted that hydrogen bonding and non-hydrogen bonding species need to be treated separately[56,57].

The use of RP-HPLC retention to calculate log *P* values has been the subject of much work. Kaliszan[1] listed over 100 references in which log *P* was calculated using either RP-HPLC or TLC retention. Using the observed linear relationship, it should therefore be possible to calculate the capacity factor of a compound from its known log *P*. However complications have arisen as the slope of the relationship between log *P* and log *k'* was dependent on the column make, the eluent composition and the type of compounds being studied. Braumann[54] has suggested that at 0% organic modifier there should be unit slope between log *k'* and log *P* as at this composition the system was very close to the octanol-water system. He also suggested that at 0% organic modifier the values obtained should be independent of the stationary phase[58]. However, it was not usually practical to use eluents containing 100% water due to the long retention times of most compounds and these values were obtained by linear extrapolation of the retention data. However, as the change in capacity factor with eluent composition has been best described using quadratic regression equations,[59,60] the linear extrapolation could result in considerable errors in the 100% water value of capacity factor. The use of quadratic equations to extrapolate the data might also result in considerable problems because of the volume of data which would be required for an accurate result.

Hanai *et al.* have used the observed linear relationship between log *k'* and log *P* to predict the retention of a variety of compounds including nitrogen containing compounds[61], phenols[62,63] and in combination with dissociation constants and Van der Waals radii to predict

the retention of acids[41,64-66]. In each case the column/eluent system was calibrated using a set of compounds structurally related to the "unknown" test compounds. The retention was then calculated by substituting the log *P* (usually calculated as the sum of Rekker[53] fragmental constants) of the "unknown" compound into the regression equation. Shalaby *et al.*[67,68] have used the measured log *P* values of nitrogen bridged compounds to calculate their retention.

Baker *et al.*[69-73] developed a prediction method in which the retention index of a compound was calculated from the measured retention index of a "parent" compound and a multiple of the substituent Hansch substituent constants (π)[48]. The method was used to calculate retention indices of a number of drug compounds including barbiturates[71], anthranilic acid analogs[69], glucuronide metabolites[72] and steroids[73].

Jinno and co-workers have used π constants to predict the retention of alkylbenzenes[74,75]. A combination of molecular connectivity index values, hydrogen donating and accepting terms (HA-HD), Van der Waals volume, Hammett constants and π values have been used to calculate retention of a number of different compounds[28,29,76,77]. Multiple regression analysis was used to determine which parameters were most significant and the contribution of each for a particular type of compound. The choice of which of the parameters to use was found to depend on the class of the compound and also the length of the bonded phase of the column. This prediction system was written to run on a microcomputer and communicate with a database containing the regression equations relating *k'* to the structural

parameters[29,30,77,78]. The different equations required for phenyl, C_2, C_8, and C_{18} columns were contained within the database. The user was required to input the type of compound either as a name, using the molecular formula or by selecting the functional groups. As the retention prediction method was based on capacity factors the regression equations were only strictly valid for a single column. The user entered information on the retention of a set of standards run on the column to calibrate the system.

A similar approach to predicting the retention of phenols was used by Li and Lee[79] in ternary eluents but unlike Jinno *et al.* the relationship derived did not use linear combinations of the parameters. The final equation involved Hansch substituent constants to the power of 4, cubed, and squared plus a multiple of the Hansch substituent constant and Hammett constant for the compound. The accuracy of the prediction ranged from +/- 2% to +/- 12% depending on the complexity of the molecule. This method not only requires the retention times at a minimum of seven eluent compositions to obtain the regression coefficients but considerable computing power to perform the stepwise multiple regression equations. It is unlikely that the derived equations are generally applicable to different classes of compounds or different columns.

1.2.3 Prediction Based on Aqueous Solubility

Hafkenscheid and Tomlinson[80-82] have reported a linear relationship between RP-HPLC retention data and aqueous solubility parameters. Acids and alcohols had to be treated separately from basic and neutral molecules[81]. Although the

RP-HPLC capacity factors were useful to calculate aqueous solubility[80-82] the use of the Hildebrand solubility parameter to calculate retention resulted in an over-estimation of the retention by 4-5 times[83]. Tijssen *et al.*[84] also suggested that aqueous solubility could be used to predict the selectivity and solubility in chromatography.

An alternative "retention-solubility parameter", R, was suggested by Jinno[85] to predict the retention of phenylthiohydantoin amino acid derivatives. The solubility of these compounds was expressed relative to that of phenylthiohydantoin lysine which had the smallest aqueous solubility. There was a linear relationship between R and log k' which was used to predict the retention of other derivatives. A separate regression equation was required at every eluent composition.

1.2.4 Prediction based on Solvatochromic Parameters

Two groups of workers have suggested the use of solvatochromic parameters for retention prediction. The solvatochromic parameters were described by Kamlet and co-workers[86] as the "solvent" effects which cause shifts in UV/VIS spectroscopy. The parameters described include a measure of the solvent-solute dipole interactions and the hydrogen bonding interactions of the solute.

Dorsey and co-workers[87-89] have suggested the use of the solvent polarity parameter, $E_T(30)$, to describe the change in retention of a variety of compounds as an alternative to organic modifier concentration. A linear relationship was observed for acetonitrile eluents but not for methanol eluents. The solvent polarity was also used to

obtain methylene group selectivity values for different series of compounds and columns.

Kamlet and co-workers[89,90] have used a number of polarity and other solvatochromic parameters to describe the retention of a number of compounds. The method was suggested to calculate either the adjusted retention time or the capacity factor using a combination of a measure of the molar volume (V_1), the solute polarity (π^*), which was related to the solute dipole moment, and the hydrogen bond acceptor strength (β). These factors were suggested as being the dominant ones in determining the retention of a solute [91,92]. It was also suggested that the results could be used for both prediction and characterisation of the stationary phase. The regression equations depended on both the eluent composition and the column although it was suggested that the signs and relative size of the different contributions are always the same. A typical equation is shown below[90]:-

$$\log k' = -(0.49 \pm 0.03) + (1.207 \pm 0.026)V_1/100 - (0.110 \pm 0.038)\pi^* - (0.764 \pm 0.085)\beta \qquad 1.7$$

However these relationships have not been used to calculate the retention of unknown compounds.

Recently the method has been extended to examine the retention of a number of classes of compounds on different stationary phases[93]. An equation similar to equation 1.7 was derived, although the coefficients differed. The equation (1.8) was used to calculate the retention for a number of substituted benzenes, polynuclear aromatic hydrocarbons and polychlorinated biphenyls.

$$\log k' = (-0.58 \pm 0.04) + (1.88 \pm 0.04)V_1/100 - (0.45 \pm 0.04)\pi^* - (1.47 \pm 0.08)\beta_m - (0.46 \pm 0.13)\alpha_m. \quad 1.8$$

where V_1 is a measure of the molar volume, π^* the solute polarity, β_m the hydrogen bonding acceptor strength of the solute and α_m the hydrogen bonding donator strength of the solute.

1.2.5 Prediction Methods using Group Contributions

In the prediction methods using group contributions, the retention of a compound is calculated as the contribution from a parent molecule plus contributions from the substituents. This approach is very similar to that used by Hansch[48] to calculate octanol-water partition coefficients described in a previous section (Section 1.2.2). The major difference between these prediction methods and those described previously was the direct use of retention parameters rather than the use of structural parameters of the molecule.

Two approaches have been used to develop retention prediction methods by this method. The first used capacity factors which have been shown to be very dependent on the chromatographic conditions used for their determination. The second used group contributions based on retention indices in which the retention of a compound was expressed relative to a set of standards. The use of retention indices in reversed-phase high performance liquid chromatography will be discussed in a later section (Section 1.3). Although the general approach was similar, the data obtained cannot be

directly compared and there were differences in the calculation methods. The two approaches will therefore be discussed separately.

a) Group Contributions using Capacity Factors

The group contribution (T) for most of these methods was defined using equation 1.9[94]:

$$T = \log (k'_y - k'_p) \qquad 1.9$$

where T was the group contribution, k'_y was the capacity factor of a substituted compound and k'_p the capacity factor of a "parent" compound. Alternative definitions of the group contribution have been used and will be discussed below. From the equation above the capacity factor of a compound (X) was calculated using the equation

$$\log k'_x = \log k'_p + \Sigma T \qquad 1.10$$

The papers differed in the choice of parent used for the initial calculation of the group contributions. These included benzene and substituted benzenes[95-98], esters[99], coumarins[100], catecholamines[94,101], 2-phenylethylamine[102], purine[103], chromone and isoflavone[104], and steroids[105]. There were complications in the use of complex parents, in that it was not always possible to derive group contributions in the absence of interactions with substituents integral to the parent structure. The values were therefore specific to that system.

Two methods for calculating the group contributions

have been suggested. Firstly, the straightforward subtraction of the values for two compounds and secondly, calculation of the group contributions by solving sets of linear equations. These latter equations express the capacity factors as the sum of the group contributions for the different substituents and were probably more readily automated.

The papers described below reported calculated group contributions and suggested that they could be used for retention prediction. However, relatively few have actually used the values for calculating capacity factors of "unknown" compounds.

Molnar and Horvath[94] proposed one of the earliest group contribution methods in which the parent was either an acid or catecholamine. *T* values were obtained for hydrogen, hydroxyl, methoxy, and amino substituents. A larger range of substituents were examined by Chen and Horvath[101] and sets of linear equations were solved to obtain values for the substituents and also for the interactions occurring between the substituents. The values were obtained on three columns and reasonable correlations between the different columns were observed, but again the values were not used for retention prediction. Horvath and co-workers[106] have adopted a similar approach to benzoquinones and hydroquinones to obtain *T* values. These were used in combination with molecular connectivity indices to calculate the capacity factors of five benzoquinones.

Tomlinson *et al.*[99], Gill *et al.*[102], and Glowniak and Bieganowski[100] calculated *T* values by subtraction. Tomlinson *et al.* examined a series of substituted alkyl benzoates with OCH_3, NO_2 or Cl in various positions on the aromatic ring[99].

The experimental data suggested that the relationship between the carbon number of the ester chain and log capacity factor was non-linear. The methylene group contribution increased with increasing chain length up to an ester chain of five carbon atoms. The group contributions were dependent on the position relative to the ester group, the eluent composition and the column but were apparently not dependent on which alkyl benzoate was parent. The derived *T* values were correlated with thermodynamic properties but were not used for prediction.

Gill *et al.* examined 30 substituted phenylethylamines to derive *T* values for methyl, hydroxyl, methoxy, amino and N-oxide in a number of positions and combinations[102]. With these compounds the group contributions were dependent on the substituted compound and its "parent" because of interactions occurring with the other substituents. The group contribution of a methyl group on an aromatic ring ($T = 0.54$) differed significantly from the values for methyl substituents on the side chain ($T = 0.36 - 0.44$).

Glowniak and Bieganowska[100] examined a set of 29 coumarins, furocoumarins and pyranocoumarins including three parent species coumarin, psoralen and xanthyletin. A number of substituents were examined either singly or in combinations. The group contributions were found to depend on both the position and the parent compound, again this was probably due to substituent interactions.

One of the few papers which has used the group contributions for prediction was by Assenza and Brown[103]. Using purine as the parent the contributions were calculated by solving sets of linear equations to obtain the values from 24 substituted compounds with a range of substituents,

methyl, oxo, amino, methylamino, dimethylamino, thio, imino and ribonucleosides in various positions. Group contributions were obtained on four columns using a single eluent composition. The values obtained differed on the different columns although there was a high correlation between two C_{18} columns. The data were then used to calculate the retention of an additional 62 purines with reported high correlation.

Bylina *et al.*[107] used the "group contribution" approach to calculate values for molecular fragments which could then be used to predict retention. Eight alkyl fragments were defined, six of which described the presence of an aromatic ring and the other two to describe the alkyl portion of a molecule. Using this method the capacity factor would be calculated from the equation

$$\ln k = \Sigma \ln k_j + \text{constant} \qquad 1.11$$

the value of the constant was taken as the log of the phase ratio of the volumes of the stationary and mobile phases. The method was used to calculate the capacity factors of polynuclear aromatic hydrocarbons but it was suggested that the values could be extended to other compounds.

Borda *et al.* examined a set of chromones and isoflavones to derive group contributions for methyl, methoxy, and hydroxyl groups[104]. However the definition of the group contribution differed from that described earlier. The group contribution was based on the "separation factor"(α) which was defined in terms of the capacity factor rather then log capacity factor ($\alpha = k'_i/k'_j$). The size of the separation factor was dependent on the other

substituents present.

A slightly different definition of the substituent contribution ($k\Gamma$) was proposed by Sadek *et al.*[95] using the equation:-

$$\delta\Gamma = -RT \ln (k'_x/k'_b) \qquad 1.12$$

where R is the gas constant, T the temperature, and k'_x and k'_b the capacity factors of the substituted benzene and benzene. The substituent contributions were derived on a number of stationary phases and it was that suggested that they could be used to characterise the stationary phases but the values were again not used for prediction.

A different approach was taken by Tsantili-Kakoulidou *et al* [96]. Log k' values were extrapolated to 0% organic modifier, to obtain log k'_w values. In an analogy with previous work on the use of the use of log k_w values to calculate log P[54,58,108,109] ,(see Section 1.2.3) the group contributions were termed π^*. A range of monosubstituted compounds were used to obtain the π^* values and 43 disubstituted benzenes were examined to investigate substituent interaction effects (T^*), using the equation:-

$$\log k'_{exp} = \log k' + T^* \qquad 1.13$$

The values were used to calculate the retention, at 0% organic modifier, of some polysubstituted compounds with a high correlation between the calculated and extrapolated retention values. Braumann and Jastorff[109] also used this definition of the group contribution to calculate the capacity factor at 0% organic for a number of cyclic nucleotides, although the aim was to predict the

quantitative structure-activity relationships not the retention.

A related method was previously described in the section on Van der Waals volumes (see Section 1.2.1 d) in which a hydrophobic group contribution was derived and used for calculating retention[42].

b) Group Contributions using Retention Indices

A number of papers have calculated group contributions based on retention indices using the equation

$$\Delta RI = (RI_m - RI_p) \qquad 1.14$$

where ΔRI is the retention index group contribution, RI_m the retention index of the substituted compound and RI_p the retention index of the parent. The papers described in this section have again used a number of parents, benzene[110], nitrobenzene[111], antioxidants[112], ergopeptides[113] and alkylbenzenes[114]. As the group contributions were often based on different retention indices scales the resulting values cannot be readily transferred from one study to another.

Popl *et al.* calculated substituent contribution to retention index based on a set of monosubstituted benzenes[110]. Substituent contributions were calculated using a single column and a single eluent composition. These were then used to calculate retention indices of polysubstituted benzenes. Significant differences between the calculated and measured values were observed if there were interactions between substituents, for example with nitrophenols,

polyphenols and compounds containing carboxylic acid groups. Good agreement was found for halogens, phenols and nitro compounds other than those mentioned previously. A set of substituted nitrobenzenes[111] was used to examine substituent interactions but the values were not used for prediction. An extension of the study also derived substituent contributions from phenolic antioxidants[112]. These were extrapolated to 0% methanol and the regression equations used for prediction at other eluent compositions. Several standards were used to calibrate the system but the use was limited to a few groups, including secondary carbon groups, tertiary carbon groups, alkyl carbons, aromatic rings and hydroxyl groups.

Magg and Ballschmiter[113] derived substituent contributions on four columns at a single eluent composition using an ergopeptide parent. For the substituents examined methyl, *i*-butyl, benzyl, ethyl, *i*-propyl and *s*-butyl the group contributions were virtually unchanged on the different columns although the actual retention indices varied considerably. It was suggested that the values could be used for retention prediction but no results were given.

Retention index substituent contributions have also been determined by Morishita *et al.*[114] for a single column and eluent composition. They measured values for methyl, hydroxyl, amino and nitro groups using benzene as the parent. Positional increments were also obtained for the pairs of substituents. These were then used to calculate the retention index of some polysubstituted benzenes.

c) Prediction Based on Interaction Indices

As an alternative to retention indices an interaction index scale has been proposed[115,116]. The interaction index was defined as a measure of the polar interaction occurring in the mobile phase and was calculated using a calibration line and five standards. It was suggested that the interaction index of a compound could be used to predict its log k' using the calibration line[115]. Jandera[117-121] proposed another prediction method which was also based on the model of interaction indices. The retention of a compound was divided into two contributions, a non-polar (n_{c_w}) and a polar contribution (q). The polar contribution was found to be different for different substituents but was independent of the parent compound while the non-polar contribution was found to be dependent on the parent. The resulting values were used to calculate the capacity factors a range of substituted benzenes[118,119] and phenylurea and triazine herbicides[121].

1.2.6 Miscellaneous Retention Prediction Methods

a) Mobile Phase Relationships

Many optimisation methods use mobile phase relationships to predict the behaviour of compounds on changing eluent compositions. These methods are interpolation or extrapolation methods rather than "prediction". However two methods have been suggested by Baty and Sharp[122] and Cooper[123] that also use the molecular structure of the compound. In these a set of calibration

standards structurally closely related to the analyte were used to obtain equations describing the change in retention with eluent composition. Using the calibration line it was possible to calculate the retention of an "unknown" compound from a single measured capacity factor. These methods are also therefore extrapolation or interpolation methods rather than prediction solely from molecular structure.

Another retention prediction method based on mobile phase relationships was recently suggested by Marengo *et al.*[124]. The aim in this paper was to calculate the retention time in gradient elution from initial isocratic data. The velocity at which a solute moved through the column was calculated from the isocratic data. This was then used to predict how the solute would behave under gradient conditions.

b) Prediction Based on Thermodynamic Parameters

Jaroniec[125,126] described the retention of solutes in terms of their thermodynamic properties such as activity coefficients and suggested that these could be used for prediction. Petrovik *et al.*[127] also suggested the use of activity coefficients for prediction of retention in terms of a functional group contribution in which the retention was expressed as the sum of the activity coefficients plus a constant. These workers also suggested a method for calculating the activity coefficient in mixed mobile phases.

c) Prediction using the Electronic Interaction Index

Lamparczyk and co-workers[128] have proposed an

electronic interaction index for describing the retention in GC, HPLC and TLC. The electronic interaction index was calculated from the ionisation potentials, molecular polarizability and dipole moments and it was suggested that the values could be useful for predicting the retention properties of solutes[129,130].

d) Computer Based Prediction Methods

A number of prediction methods have been based on computer systems. These will be discussed more fully in a later section (Chapter 8) but have included the database system of Jinno and co-workers[29,77,78] and the many optimisation systems.

1.3 USE OF RETENTION INDICES IN RP-HPLC

In RP-HPLC the commonest method of reporting retention is the capacity factor (k') calculated according to the equation $k' = (t_R - t_o)/ t_o$. Although capacity factors are independent of the column dimensions and the flowrate they have been shown to be very dependent on the value of t_o, the temperature and the column packing material. This makes the comparison of data from different sources very difficult. In attempting to find a more robust method of reporting retention various workers have suggested the use of retention indices. Most retention index scales are based on the linear relationship, for a homologous series, between the carbon number and log capacity factor. HPLC retention indices are therefore analogous to Kovats indices, which have frequently been used to standardise retention in gas-

liquid chromatography[131], in which the retention of a compound is expressed relative to the retention of n-alkane standards. One of the problems in LC is the choice of suitable standards, n-alkanes are not generally applicable due to the problems of the lack of a UV chromophore and the limited polarity range. A successful standard homologous series should have a reasonably strong UV chromophore, preferably at the most commonly used wavelength of 254 nm. The standards should be widely available and should cover the polarity range of typical analytes. Under the experimental conditions the standards should also not be ionised. Many homologous series have been suggested as suitable standards, these have been described in a recent review by Smith[132] and include n-alkanes, n-alkylbenzenes, alkan-2-ones, alkylarylketones, esters and polycyclic hydrocarbons. Of these the alkan-2-one scale proposed by Baker[133] and the alkylarylketone scale proposed by Smith[134] have been most widely used.

An additional scale has recently been proposed by Bogusz[135] using the n-nitroalkanes as standards. Although these last compounds have a reasonable chromophore at about 230nm their absorption is very small at 254 nm. Bogusz suggested that as the early members of the series have short retention times, they could therefore be used as an alternative to the alkan-2-one scale to determine the retention indices of poorly retained drug species. These compounds often elute before the first member of the alkylarylketone scale and the curve would have to be ~~to be~~ extrapolated.

Bogusz[136-138] has also published work describing the use of "corrected "retention indices. In these papers the

retention indices, based on alkylarylketones, were corrected to account for differences between columns and laboratories. These differences arise from the use of different columns and/or slightly different conditions and eluents. Although retention indices are considerably less sensitive to the exact experimental conditions they do still show some dependence.

In his review which covered the literature up to mid 1987 Smith[132] identified areas of use of retention indices as the identification of "unknown" compounds, retention prediction and structure-activity relationships, and characterisation of selectivity and stationary phases in a manner similar to McReynolds constants and Rohrschneider constants used to characterise GLC columns.

Since publication of the review two compilations of retention index values have been reported, in both of these compounds are identified by a combination of their retention indices, using alkylarylketones standards, and UV/VIS spectrum. The first of these by Hill and Langner[139] was a database of 157 drug compounds measured under acidic conditions and 144 drug compounds under basic conditions. The second by Frisvad and Thrane[140] contained data on mycotoxins and fungal metabolites.

Other recent papers which have appeared include those of Mockel[44-47] using the n-alkane scale. The retention properties of various compounds were related to structural and physical properties. Although it was suggested that this might provide a method of predicting retention this was not actually done. Recent work by Dimov[141] on retention indices has also suggested a method to calculate the n-alkane retention index of alkylbenzenes on the basis of the carbon

number and branching terms.

1.3.1 Retention Indices in Retention Prediction

Of the many retention prediction systems which have been discussed in this chapter the majority have been based on capacity factors with relatively few using retention indices. One of the major problems of basing the retention prediction on capacity factors was the high degree of dependence on the column, temperature and eluent composition which may limit the applicability of the scheme or require recalibration for every column used. Retention indices have been shown to be robust to small changes in the experimental conditions although they do change over wider eluent ranges. Retention indices were also sensitive to selectivity changes between columns but again the change was considerably less than in the capacity factor[142]. Retention prediction schemes based on retention indices should therefore prove more robust than schemes based on capacity factors. Despite these apparent advantages very few retention prediction schemes have been based on retention indices. Most of these have already been discussed in the different sections of this chapter.

Among the prediction schemes based on retention indices a variety of scales have been used. The alkan-2-one retention index scale was used by Baker[69-73], using the retention index of a parent plus a multiple of the octanol-water partition coefficient (see Section 1.2.3). Using the same retention index scale Shalaby[67] calculated retention index scales directly from the octanol-water partition coefficient of the "unknown" compound. The n-alkane scale

was used by Mockel *et al.*[44-47] in describing the relationship between retention and molecular surface area (Section 1.2.1 e).

Several workers have suggested the use of a group contribution approach using different scales (see Section 1.2.5). The n-alkane scale was used by Morishita[114]. Popl[112]. used the PAH scale and the alkan-2-one scale has been used by Magg and Ballschmiter[113].

Finally Dimov[140] has recently described a method of calculating the retention indices, on the n-alkane scale, of alkanes and alkylbenzenes based on the chain length and branching of alkylbenzenes.

1.4 INTRODUCTION TO CURRENT WORK

The aim of the current study was to develop a method to predict the retention index of a compound from its molecular structure. The approach taken was a group contribution (see Section 1.2.6) approach in which the retention index has been calculated from the contributions of the parent, the substituents and interactions between substituents using the equation

$$RI = PI + \Sigma SI_{Ar-x} + \Sigma SI_{Al-x} + SI_{R} + \Sigma II_{Y-Z} \qquad 1.15$$

where PI = Retention index of a parent

SI_{Ar-x} = Substituent index of the aromatic substituent

SI_{Al-x} = Substituent index of aliphatic substituents

SI_{R} = Substituent index of saturated alkyl chain

II_{Y-Z} = Interaction indices to account for interactions between substituents

The retention index scale used was based on the linear relationship between log capacity factor and carbon number of the alkylarylketones.

Each term in the above equation was related to the organic modifier concentration of the eluent using a quadratic equation. This enabled the retention index to be calculated at any eluent composition and therefore by calculating the retention of two compounds their separation can be optimised.

The first stage in developing the prediction scheme was to obtain numerical values for each term in the equation. Each of these terms was calculated from the retention index increment which was the difference between the experimental retention indices of the substituted compound and the unsubstituted parent at the same eluent composition.

For example, the retention index increment used to calculate the substituent index, of an aromatic substituent X would be calculated using the equation

$$\delta RI = RI_{Ph-X} - RI_{Ph-H} \qquad 1.16$$

Where the substituent was a single functional group the retention index increments can then be used to determine regression equation describing the change with eluent composition. This equation can then be used to calculate the substituent index at any eluent composition. A single parent, benzene, was used throughout the study and substituent indices were calculated for 14 aromatic substituents and 12 substituents on aliphatic side chains. Substituent indices have been obtained for a range of substituents using model compounds which were

monosubstituted benzenes or alkylbenzenes.

The interaction indices were calculated from compounds in which more than one substituent was present (e.g. XC_6H_4Y). The interaction increments (δI) were calculated from the difference between the measured retention index of this compound and the sum of the parent retention index and substituent indices for X and Y.

$$\delta I = RI_{XPhY} - (PI + SI_X + SI_Y) \qquad 1.17$$

As with the substituent indices these have been obtained using a number of model disubstituted benzenes over a range of eluent compositions.

It was important that the increments were reproducible and as robust as possible. To ensure this and that the observed changes in retention index were due to changes in the substituents or eluent selectivity a single batch of column packing material was used throughout the study and the experimental conditions were carefully controlled.

Having obtained the set of equations the next stage was to develop a method to calculate the retention index of a compound from its molecular structure. This has been done using an expert system to access the database of regression equations. The success of the approach in calculating retention indices has been tested using a set of test compounds.

A commercial prediction / optimisation system was also examined by comparison with the experimental data obtained during this study.

CHAPTER 2

EXPERIMENTAL

2.1 CHEMICALS

a) Mobile Phase Components

Methanol, acetonitrile and tetrahydrofuran were HPLC grade (FSA Laboratory Supplies, Loughborough).
Sodium dihydrogen orthophosphate and disodium hydrogen orthophosphate were AR grade (FSA Laboratory Supplies, Loughborough).

b) Retention Index Standards

Acetophenone, propiophenone, butyrophenone, valerophenone, hexanophenone and heptanophenone were from various sources.

c) Model and Test Compounds

The model compounds, except phenylacetamide and 3-phenylpropionamide, and the test compounds were laboratory grade purchased from various sources.

3-phenylpropionamide and phenylacetamide were prepared from the corresponding acid chlorides by reaction with ammonia.

d) Void Volume Marker

Sodium nitrate AR grade (FSA Laboratory Supplies,

Loughborough).

2.2 HPLC EQUIPMENT

All experimental work was performed using a HPLC system consisting of a Pye-Unicam PU 4010 pump and a Pye Unicam PU 4025 UV detector set at 254 nm. Separations were carried out using a 100 x 5 mm I.D. column slurry packed in the laboratory with 5 µm Spherisorb ODS2 (batch 23/151, Phase Separations, Queensferry, U.K.). Injections were made using a Rheodyne 7125 injection valve fitted with a 20 µl loop. The column was maintained at 30°C by circulating water, from a water bath thermostatted to 30°C, through a glass jacket surrounding it. Retention times were recorded using a Shimadzu Chromatopac CR-3A integrator.

2.3 COMPUTING

Calculations of retention indices were performed on an APPLE II computer and later on an OPUS PC II personal computer, Opus Technology, using a least squares program written in BASIC. Regression curves were fitted using a curve fitting program, Curve Fitter, Interactive Microware Inc., on the APPLE II. Retention prediction was performed using an expert system shell VP-Expert, Paperback Software, and a spreadsheet package VP-Planner, Paperback Software, on the OPUS PC II. An optimisation package ,Drylab I, LC Resources inc. was also used.

The OPUS PC II had 1024K RAM, dual 5 1/4" 360K disk

drives and monochrome monitor. It was fitted with a Hercules type monochrome graphics card and was attached to a printer (Panasonic KX-P1081).

2.4 EXPERIMENTAL PROCEDURE

2.4.1 Mobile Phase Preparation

The aqueous phase was a pH 7 buffer prepared from 1.5 g of sodium dihydrogen orthophosphate and 1.3 g of disodium hydrogen orthophosphate dissolved in 1 l of di-distilled or deionised and scrubbed water. The buffer was used undiluted except at 90% organic modifier where it was necessary to dilute the buffer 10 fold to prevent precipitation of the buffer salts.

Mobile phases containing 40 to 90% methanol and 30 to 90% acetonitrile at 10% increments were prepared by volume from the measured volume of the organic modifier and adding a measured volume of the aqueous buffer solution. The same mobile phase compositions were used on different occasions.

The eluents were degassed under vacuum using a water pump before use.

2.4.2 Solute Preparation

The retention index and solute solutions were prepared by dissolving sufficient of the compound in 10 ml of mobile phase to enable detection using the UV detector set at 0.08 aufs.

2.4.3 Void Volume Solution

The void volume marker solution was prepared by dissolving sodium nitrate in water at a concentration of 6 mg ml^{-1}. The void volume was taken as the peak maximum as recorded by the integrator.

2.4.4 Column Testing

During the study four columns packed in the laboratory from the same batch of packing material were used in sequence. A column was repacked with fresh stationary phase when the peak shape deteriorated.

The columns were packed using an upward slurry packing method at a pressure of 4000 - 6000 psi. The packing material (1.8g) was suspended in 5ml of propan-2-ol and packed in propan-2-ol. Methanol-water (50:50) was used to condition the column.

After packing the columns were tested for efficiency and peak symmetry using an eluent containing 70:30 methanol - water and a test solution containing benzamide, acetophenone, benzophenone and biphenyl. The efficiency was measured on the biphenyl peak and was in the range 5000 - 6000 for each 100 mm column.

To ensure that the retentive capacity of the different columns was comparable each column was tested using a set of model compounds (see later Chapter 3, Section 3.4) at a single eluent composition (MeOH-buffer 60:40). On a daily basis the column performance was monitored using three model compounds, phenol, benzene and toluene.

2.4.5 Experimental Method

The column was allowed to equilibrate with mobile phase for at least 1 hour at 1 ml min^{-1} flowrate with the effluent going to waste. After this the eluent was recycled to the solvent reservoir.

The flowrate was altered as necessary between 0.5 and 3 ml min^{-1}, depending on the eluent composition, to maintain reasonable retention times.

An injection (10 µl) of the retention index standards, acetophenone to heptanophenone, as a mixture was made followed by an injection of each model and test compound individually and the void volume marker. This sequence was repeated three times for each eluent composition. Where possible the three sets were completed on a single day however where this was not possible a single batch of the mobile phase was used over a two day period to complete the set.

2.5 CALCULATIONS

Mean retention times were calculated as the arithmetic mean of the three retention times measured for a solute and this value was used for all subsequent calculations.

2.5.1 Capacity Factors

Capacity factors were calculated according to the equation

$$k' = (t_r - t_o) / t_o \qquad 2.1$$

and expressed to two decimal places.

2.5.2 Retention Indices

Retention Indices were calculated as described by Smith[134]. A linear least squares regression equation was obtained for log k' *vs.* 100 x carbon number for the alkylarylketone standards run as part of the same sequence of injections. The retention indices are obtained for the test solutes by substitution of their capacity factors into the regression equation. A computer program was written in BASIC to carry out the calculation initially on an APPLE II computer and later on an OPUS II.

2.5.3 Parent indices

The parent indices were calculated from the quadratic regression equations relating the change in the retention index of the parent (benzene) to the methanol or acetonitrile proportion of the eluent over the range 30 - 80% MeCN or 40 - 80% MeOH.

2.5.4 Substituent Indices

Retention index increments for a substituent were calculated from the equation

$$\delta RI = RI_x - PI \qquad 2.2$$

where RI_x was the retention index of a substituted compound and PI was the calculated parent index. Substituent indices (SI) were calculated from quadratic regression equations relating the change in the retention index increments to eluent composition (over the ranges 30 - 80% acetonitrile and 40 - 80% methanol).

2.5.5 Interaction Increments

Interactions between substituents were measured using disubstituted compounds. The interaction increments were calculated from the difference between the measured retention index of these compounds and the sum of the parent index and the substituent indices of the substituents present.

$$\delta I = RI_{PhxY} - (PI + SI_x + SI_Y) \qquad 2.3$$

2.5.6 Retention Prediction

Retention indices of "unknown" compounds were calculated using the expert system VP-Expert on an OPUS II personal computer as will be discussed in a later section.

CHAPTER 3

REPRODUCIBILITY AND ROBUSTNESS OF CAPACITY FACTORS AND RETENTION INDICES

The aim of the study was to develop a method of calculating the retention of a compound based on its molecular structure. The proposed method to do this was based on the summation of the contributions to the retention index from the parent structure, the substituents and any interactions between the substituents. The retention index, at a particular eluent composition, can therefore be calculated from the equation introduced earlier:-

$$RI = PI + \Sigma SI_{Ar-x} + \Sigma SI_{Al-x} + SI_{R} + \Sigma II_{Y-Z} \qquad 3.1$$

where RI = retention index of compound

PI = retention index value of a parent compound

SI_{Ar-x} = substituent index values for aromatic substituents

SI_{Al-x} = substituent index values for aliphatic substituents

SI_{R} = retention index contribution of saturated aliphatic carbons

II_{Y-Z} = interaction index values due to the interactions between the substituents

Schoenmakers *et al.*[59,60] have described the relationship between retention expressed as log k' (which is directly related to the retention indices) and the organic

modifier concentration as quadratic within the eluent range 10 - 90 % organic modifier. In eluents containing less than 10% organic an additional term was required to describe the behaviour. However, it was suggested that for many compounds, over a more limited range the quadratic term was not significant and a linear relationship could be valid. This assumption has been used by several workers in prediction and optimisation schemes. It has also been used to obtain k' values at 0% organic modifier by extrapolation, these have been used in the calculation of octanol-water partition coefficients (Section 1.2.3) and were also used by Testa and co-workers[96] in their retention prediction method. In the present study a quadratic relationship was generally assumed for the change in retention index increments and related functions with eluent composition, unless the change with composition was very small.

The initial stages of the study have involved the accumulation of a database of substituent indices and interaction indices based on experimental data from model compounds over the eluent ranges 30-80% acetonitrile and 40-80% methanol. The changes in the parent index, the substituent indices and the interaction indices with organic modifier concentration are described by quadratic regression equations. The regression equations can then used to predict the retention indices of other compounds. The experimental determination of these values will be discussed in the following chapters.

The accuracy of any retention prediction using the above equation will be dependent on the reproducibility and robustness of the index values to small changes in conditions. The values need to be reproducible over the

period of the study and also it must be possible to reproduce the values in other laboratories and on other columns. To ensure that any changes observed were not due to changes in the experimental conditions a number of parameters were carefully controlled throughout the study. These included temperature, the mobile phase and the column. A description of the methods used throughout the study is given in this Chapter (Section 3.1). The effect of the inevitable small variations in the experimental conditions on the capacity factor and retention indices have been investigated. The results have been used to establish the uncertainty which would be expected in any experimental or predicted retention index value.

3.1 METHOD PROTOCOLS

To ensure that the changes in the measured index values were a result of changes in the substituent and not due to changes in the separation conditions a number of experimental conditions were controlled. By following a constant experimental procedure it should be possible to reproduce the values both on the same equipment and on other systems.

a) Temperature

Capacity factors (k') have been shown to be inversely proportional to the absolute temperature[9,143] and temperature has also been shown to effect the retention indices[144] of some compounds. Despite these observations

many workers still work at ambient temperature but in this laboratory the ambient temperature varied considerably over even a short period. The temperature of the column was therefore maintained at 30°C by circulating water from a thermostatted water bath through a glass jacket surrounding the column. This temperature was above the usual maximum ambient temperature of the laboratory and could therefore be readily maintained. Previous work using retention indices on the alkylarylketone scale had also used this as the standard temperature[144-149].

b) Buffer pH

The retentions of ionisable compounds have been shown to be dependent on the pH and ionic strength of the mobile phase as the extent of ionisation varies[144]. The aqueous component of the mobile phase was therefore standardised as a pH 7 buffer prepared by weight from sodium dihydrogen orthophosphate and disodium hydrogen orthophosphate. Although using a pH 7 buffer ensured that the majority of compounds studied would not be ionised, it was not possible to examine carboxylic or sulphonic acids which would be fully ionised at this pH. The results also suggested that the primary aliphatic amines might be partially protonated at this pH. The ionic strength of the aqueous phase was kept constant by preparing the buffer by weight. However, it was necessary to dilute the buffer tenfold for eluents containing 90% MeCN or 90% MeOH to prevent precipitation of the buffer salt components. The pH of the buffer was not altered by this action and the effect of changing the buffer concentration of the aqueous phase on the retention of a

selection of compounds has been studied (see Section 3.3).

c) Stationary Phase

Changes in capacity factor and retention index have been reported where different makes and batches of column packing material have been used[142]. To ensure that any changes in different index values throughout the study were not caused by changes in the packing material a single batch of Spherisorb ODS2 packing material was used throughout the study. Four columns, packed in the laboratory, with the same batch of packing material were used in turn. A column was replaced when the peak shape deteriorated.

d) Experimental Procedure

The experimental procedure followed throughout the study was kept constant. The injection procedure for a set of runs was an injection of the alkylarylketones (acetophenone to heptanophenone) as a mixture followed by the individual test and model compounds and finally the column void volume marker solution. This procedure was repeated three times for each set of compounds. Where possible the three runs were completed on a single day, however with eluents containing low organic modifier concentrations this was not possible. The three runs were completed using a single batch of eluent which was recycled. The column void volume was measured from the retention time of the UV maximum absorption peak, as recorded by the integrator, of a constant injection volume (10 µl) and concentration (6 mg ml^{-1}) of an aqueous solution of sodium

nitrate. However the retention time of this compound was very short and small variations in the measured retention time could significantly alter the capacity factors. Alternative methods of obtaining a value of the column void volume by calculation were therefore investigated (see Section 3.2).

The flowrate was varied between 0.5 ml min^{-1} and 3 ml min^{-1} depending on the eluent composition to maintain reasonable retention times for the model compounds.

e) Calculations

Although the integrator recorded retention times to 0.001 minute as the injections were all made manually the accuracy cannot be expressed to this precision. Capacity factors were therefore calculated to two decimal places and retention indices were expressed as integers.

The capacity factors were calculated from the arithmetic mean of the three retention times (t_R) of the solutes. The retention indices were calculated as described previously by Smith[1,3,4] from a linear regression line describing the relationship between log capacity factor and carbon number of the alkylarylketones, the correlations were usually high (see Section 3.5). Retention indices of the model compounds can then be calculated by substitution into the equation. In each eluent the retention indices were calculated using the capacity factors of the alkylarylketones included within the same set of injections as the model compounds.

Despite following the experimental procedure described

above and controlling as many factors as feasible experimental errors cannot be totally eliminated. This will be reflected in the uncertainty in any determined value and therefore in the accuracy which could be expected from the predicted retention indices. To examine the success of the precautions taken in producing reproducible retentions and to determine the uncertainty in an individual retention index several investigations of the reproducibility were undertaken. These have included a study of the short term reproducibility of the measurements on a single day and the long term reproducibility over the period of the study. The linearity of the log capacity factor *vs.* carbon number (C_{no}) relationship for the alkylarylketones, which was used to calculate the retention indices was also investigated.

3.2 THE DETERMINATION OF COLUMN VOID VOLUME AND THE EFFECT ON CAPACITY FACTORS AND RETENTION INDEX

Capacity factors were calculated from the equation $k' = (t_R - t_o)/ t_o$, they were therefore dependent on the values used for the column void volume. It was therefore necessary to use a reproducible measurement of the column void volume. Despite the importance of the column void volume there is no agreed standard method for determining its value. The suggested methods of determining the column void volume have been the subject of considerable work which has been discussed recently by Smith *et al.*[150] and Djerki and Laub[151,152]. Various methods have been suggested including the injection of unretained species, the injection of mobile phase components or deuterated mobile phase components, the

volume of the "accessible" pore space and calculation methods. Smith[150] recommended the use of an injection of uracil, however, other workers[153-155] have recommended using the retention time of an injection of a fixed volume and concentration of a sodium nitrate solution.

Although capacity factors were very dependent on the column void volume Smith[142] reported that retention indices, as an interpolated scale, were virtually independent of the column void volume value. In previous studies on the use of retention indices, based on the alkylarylketones, an injection of sodium nitrate has been used to determine the column void volume, this method was also used in this study. The dependence of the retention of sodium nitrate on the eluent composition and concentration injected has been studied below. The reproducibility of the measured column void volume has also been examined.

Because of the uncertainty in the measurement of the retention time of the unretained species it has been suggested that calculations could be used to estimate the column void volume. Two possible calculation methods were examined and will be discussed below.

The effect of using different column void volume values on the capacity factor and retention index of a selection of test compounds has been examined (Section 3.2.5).

3.2.1 Investigation of the Dependence of the Retention of Sodium Nitrate on its Concentration

Several workers have studied the retention characteristics of inorganic salts, including potassium nitrate, potassium dichromate, sodium nitrite and sodium

nitrate which have been used for the determination of the column void volume[150,154]. Sodium nitrate has probably been used most frequently as it has a strong UV absorption at 254 nm. Jinno *et al.*[155], and Hennion and Rosset[153] have found that in methanol and acetonitrile eluents the retention depended on the injection concentration and injection volume[155]. However, Wells *et al.*[154] found that provided the concentration was greater than 3×10^{-6} M the retention of sodium nitrate was independent of the injection concentration.

In the present study the dependence of the retention time of sodium nitrate on the injection concentration was investigated using 10 µl injections of aqueous sodium nitrate solutions containing 1 to 24 mg ml^{-1} $NaNO_3$. These concentrations correspond to molar concentrations of 0.01 to 0.28 M $NaNO_3$. For all the injections the retention time was taken as the maximum of the UV absorption peak as recorded by the integrator and the mean of three injections were calculated (Table 3.1). The study was carried out in two eluents, methanol-buffer (70:30) and acetonitrile-buffer (70:30), at a constant flowrate of 1 ml min^{-1}.

In both eluents the retention time of sodium nitrate was dependent on the injection concentration (in methanol the retention times ranged from 0.843 minutes for a 1 mg ml^{-1} injection to 0.922 for a 24 mg ml^{-1} injection, in acetonitrile 0.749 for a 1 mg ml^{-1} injection to 0.818 for a 24 mg ml^{-1} injection). This agreed with the findings of Jinno *et al.*[155] and Hennion and Rosset[153]. Consequently for all subsequent work a constant injection volume, 10 µl, of a constant concentration, 6 mg ml^{-1}, of sodium nitrate was used to determine the column void volume.

3.2.2 Investigation of the Dependence of the Retention Time of Sodium Nitrate on the Eluent Composition

The dependence of the retention time of sodium nitrate on the ionic strength of the aqueous phase at a constant organic modifier concentration and the concentration of organic modifier in the mobile phase have also been examined. The retention time of sodium nitrate, 0.863 minutes in methanol-water (70:30) to 0.875 minutes in methanol-buffer 0.02 M, was almost independent of the ionic strength of the aqueous phase (Table 3.2). The changes being within the range of experimental variation which were measured for MeOH-buffer 70:30 and MeCN-buffer 70:30 in the study (see Section 3.2.4, Table 3.9).

However, the retention was dependent on the proportion of organic modifier in the eluent. There was a decrease in retention time with increasing modifier concentration (0.880 minutes in 30% acetonitrile to 0.790 in 80% and 0.942 minutes in 40% methanol to 0.883 in 80% methanol) (Table 3.3). There was a marked change with the decrease in buffer concentration particularly with 90% acetonitrile (0.834 minutes) (Table 3.3).

Table 3.1: Mean retention times of sodium nitrate with different injection concentrations

Mean of triplicate 10 µl injections;
modifier-buffer (pH 7) 70:30, flowrate 1 ml min^{-1}.

Organic modifier	Retention time (min) Sodium nitrate concentration (mg ml^{-1})					
	1	3	6	12	18	24
methanol	0.843	0.847	0.856	0.897	0.917	0.922
acetonitrile	0.749	0.768	0.778	0.793	0.803	0.818

Table 3.2: Retention times of sodium nitrate with different concentrations of buffer as aqueous phase

Mean of triplicate 10 µl injections; modifier-buffer (pH 7) 70:30, flowrate 1 ml min^{-1}.

Organic modifier	Retention time (min) Buffer concentration (M)					
	0.000	0.001	0.002	0.005	0.01	0.02
methanol	0.863	0.869	0.850	0.846	0.901	0.875
acetonitrile	0.763	0.766	0.762	0.773	0.774	0.776

Table 3.3: Retention times of sodium nitrate with different organic modifier proportions

Mean of triplicate 10 µl injections, flowrate 1 ml min^{-1}.

Modifier-buffer (pH 7) v/v	Retention time (min) Modifier	
	Methanol	Acetonitrile
30:70	-	0.880
40:60	0.942	0.844
50:50	0.911	0.782
60:40	0.884	0.769
70:30	0.866	0.782
80:20	0.833	0.790
90:10*	0.826	0.834

* more dilute buffer

3.2.3 Calculation Methods for the Determination of Column Void Volume

The problem of accurately measuring the retention time of an unretained solute and the dependence of the measured retention time on the injection concentration have led to various workers suggesting alternative calculation methods, to obtain a value for the void volume. The calculation methods are based on the linear relationship between log k' and C_{no} for a homologous series[150,151,156-161]. It has been suggested that these estimated t_o values should be more reproducible than measured values. Although in most studies the use of a homologous series to determine the column void volume would be a disadvantage as it would require the measurement of additional retention times of a homologous series, this is not true if retention indices are used as the retention times of a series of homologues would already be known. It would also be possible to incorporate a routine to calculate the column void volume into the program used to calculate retention indices.

The choice of the members of the homologous series to include may be important. Van Tulder[156] recommended that the first two members of a homologous series should not be used due to possible non-linearity. A discontinuity in the log k' vs. C_{no} curve has been reported where the carbon chain length was approximately equal to the length of the bonded phase[162], in this case C_{18}. The largest member of a homologous series should therefore be less than the length of the bonded stationary phase. Two calculation methods have been examined, the first a method proposed by Berendson *et al.*[161] and the second a procedure developed by Smith[146].

Both of these calculation methods were previously examined by Smith and Garside studying retention indices on polymer columns[146]. Three lengths of the homologous series of alkylarylketones (acetophenone to heptanophenone, propiophenone to heptanophenone, and valerophenone to heptanophenone) were used to calculate column void volume values.

a) Calculation of Column Void Volume using the Berendson Method

Berendson *et al.*[161] derived an equation relating the retention times of adjacent homologues to a calculated column void volume using the equation

$$t_{R,n+1} = At_{R,n} - (A-1)t_o \qquad 3.2$$

Where t_R was the retention time of the nth or $(n+1)^{th}$ member of the homologous series. By using a least squares regression equation for $t_{R,n+1}$ against $t_{R,n}$ the slope (A) and intercept $[(A-1)t_o]$ could obtained and used to calculate t_o. This method was applied to the retention times (Table 3.4) of alkylarylketones determined in a range of eluents. The retention times of the alkylarylketones used for calculation were typical values collected on a single column.

For all the eluents and sets of homologues used the correlations for the relationship were high (0.9990 to 0.9999). The calculated t_o values (Table 3.5) were dependent on the set of homologues used and were generally higher than the values measured using sodium nitrate, for example in

methanol-buffer 50:50 the retention time of sodium nitrate was 0.678 minutes and the calculated values ranged from 1.558 to 1.990 minutes. In many cases the values were so high that they were greater than the retention times of poorly retained model compounds and it would therefore be impossible to calculate capacity factors and retention indices for these compounds. In particular the column void volume value derived from the series butyrophenone to heptanophenone in 40% methanol was clearly erroneous at 5.040 minutes as it was greater than the retention time of acetophenone (3.132 minutes). Similar high values and results have been found by Smith and Garside[146] in a study of retention indices on polymer columns.

b) Calculation of Column Void Volume using an Iterative Method

In this method an estimated value of the column void volume equal to half the shortest retention time was chosen and used in a least squares regression equation for log k' vs. C_{no}. The estimated t_o value was then changed systematically and the regression repeated until further iterations did not improve the linearity within a defined interval. The retention times and sets of alkylarylketones used previously (Table 3.4) for the calculation by the Berendson method were used to calculate t_o values (Table 3.6). The correlations observed were again high ranging from 0.9990 to 1.0000. As with the values calculated using the Berendson method the t_o values were generally larger than the corresponding measured values and the often similar to those calculated previously. Again the calculated values

Table 3.4: Retention times of alkylarylketones used for calculation of column void volume

Mean of triplicate 10 µl injections.

Carbon number of alkylaryl-ketone	Retention time (min) Modifier proportion (%) 30	40	50	60	70	80	90
methanol-buffer							
C_8	--	3.132	2.251	2.317	1.664	1.357	2.368
C_9	--	6.610	3.883	3.463	2.159	1.567	2.535
C_{10}	--	14.035	6.838	5.297	2.843	1.813	2.715
C_{11}	--	32.955	13.348	8.951	4.053	2.206	2.986
C_{12}	--	82.444	27.530	15.645	6.119	2.796	3.350
C_{13}	--	215.505	58.470	29.529	9.622	3.688	3.849
acetonitrile-buffer							
C_8	2.792	2.173	2.179	1.679	1.394	1.238	2.371
C_9	5.927	3.791	3.224	2.204	1.680	1.395	2.550
C_{10}	12.074	6.447	4.706	2.885	2.026	1.577	2.748
C_{11}	26.078	11.550	7.201	3.922	2.518	1.827	3.013
C_{12}	58.424	21.405	11.460	5.577	3.301	2.186	3.371
C_{13}	131.769	40.126	18.662	8.172	4.448	2.700	3.854

Table 3.5: Column void volume values calculated using the Berendson method

Calculated from retention times of alkylarylketones in Table 3.4..

Series used for calculation	Calculated column void volume (min) Modifier proportion (%) 30	40	50	60	70	80	90
methanol-buffer							
sodium nitrate	--	0.488	0.678	0.851	0.912	0.826	1.784
C_8 - C_{13}	--	2.408	1.558	1.234	1.149	0.970	1.922
C_9 - C_{13}	--	3.137	1.871	1.403	1.257	1.049	2.086
C_{10} - C_{13}	--	5.040	1.990	1.329	1.197	1.014	2.022
acetonitrile-buffer							
sodium nitrate	0.424	0.531	0.764	0.774	0.765	0.770	1.698
C_8 - C_{13}	0.783	0.642	0.880	0.872	0.887	0.894	1.872
C_9 - C_{13}	1.052	0.838	1.099	1.034	1.003	0.975	2.005
C_{10} - C_{13}	0.974	0.783	1.108	0.996	1.030	1.000	1.990

Table 3.6: Column void volumes values calculated using an iterative process

Calculated from retention times of alkylarylketones in Table 3.4.

Series used for calculation	Calculated column void volume (min) Modifier proportion (%) 30	40	50	60	70	80	90
methanol - buffer							
sodium nitrate	--	0.488	0.678	0.851	0.912	0.826	0.892
C_8 - C_{13}	--	0.754	1.088	1.080	1.091	0.944	0.900
C_9 - C_{13}	--	1.146	1.723	1.453	1.305	1.051	0.985
C_{10} - C_{13}	--	6.378	1.789	1.321	1.149	1.026	1.015
acetonitrile - buffer							
sodium nitrate	0.424	0.531	0.764	0.774	0.765	0.770	1.698
C_8 - C_{13}	0.447	0.492	0.698	0.774	0.839	0.832	1.760
C_9 - C_{13}	1.192	0.903	1.121	1.016	1.011	0.970	1.896
C_{10} - C_{13}	1.214	0.844	1.121	1.071	1.124	1.023	1.986

were dependent on both the eluent composition and the set of homologues.

The calculated t_o values did not give consistent results. The differences between the values calculated using the different sets of homologues suggests that the assumption of linearity may not be entirely true but dependent on the column void volume value used. The high calculated values would present problems with poorly retained species which might elute before the void volume and therefore give negative capacity factors. Overall neither of these methods offered any advantage over the use of the retention time of a standard sodium nitrate solution. Therefore for the rest of the study the retention time of 10 µl sodium nitrate (6 mg ml^{-1}) was used as the void volume marker. At this concentration the retention time was less than the retention time of the least retained model compound and the peak could be detected spectroscopically from a 10 µl injection.

3.2.4 Reproducibility of Measured Retention Time of Sodium Nitrate

The overall study has extended over approximately two years using four columns packed from the same batch of packing material. The variation of the measured column void volume over this period has been investigated, for each eluent this represents between twenty and thirty individual determinations.

The retention times of sodium nitrate varied by up to

20% from the mean retention times (Table 3.7). A change of only 2 seconds in the measured retention time would correspond to a change of 2 to 7% depending on the flowrate and eluent composition.

3.2.5 Effect of Changes in the Column Void Volume Value on Capacity Factors and Retention Indices

As measured values of the retention time of sodium nitrate varied by up to 20% from the mean value (Section 3.2.4), the effects of using t_o values of ±5% and ±15% from the mean, on the calculated capacity factors and retention indices, have been examined. Using the typical retention times of a set of alkylarylketones and a variety of compounds eluting before and after acetophenone the capacity factors were recalculated using the different t_o values (Table 3.8 and 3.9). Two methanol and two acetonitrile containing eluents have been examined, to check whether capacity factors and retention indices were more sensitive to changes in column void volume when the retention time was small.

The calculated capacity factors (Tables 3.8 and 3.9) were very dependent on the column void volume value used in their calculation. The results show the sensitivity of capacity factors to column void volume and therefore the low reliability of capacity factors due to the difficulties of measuring t_o.

These calculated capacity factors have been used to calculate retention indices (Tables 3.10 and 3.11) using the linear relationship between log k' and 100 x C_{no} for the

Table 3.7: Reproducibility of retention time of sodium nitrate

Mean of 20-30 10 µl injections of $NaNO_3$ (6 mg ml^{-1}).

Eluent	Flowrate ($ml\ min^{-1}$)	Retention time (min) Mean	S.D.	Max.	Min.
methanol-buffer (pH7)					
40:60	2.0	0.401	0.044	0.481	0.367
50:50	1.5	0.607	0.091	0.678	0.586
60:40	1.0	0.884	0.022	0.942	0.853
70:30	1.0	0.866	0.029	0.912	0.817
80:20	1.0	0.833	0.020	0.857	0.798
90:10	0.5	1.652	0.073	1.673	1.598
acetonitrile-buffer(pH7)					
30:70	2.0	0.440	0.047	0.573	0.382
40:60	2.0	0.400	0.026	0.428	0.350
50:50	2.0	0.393	0.008	0.404	0.382
60:40	1.0	0.769	0.022	0.826	0.745
70:30	1.0	0.784	0.016	0.820	0.765
80:20	0.5	1.580	0.060	1.698	1.513
90:10	0.5	1.669	0.050	1.698	1.618

Table 3.8: Effect of changes in the value used for column void volume on calculated capacity factors in methanol eluents

Compound	Calculated capacity factor Column void volume value -15%	-5%	Mean	+5%	+15%
methanol-buffer 40:60					
t_o (min)	0.341	0.381	0.401	0.421	0.461
acetophenone	9.29	8.21	7.75	7.34	6.61
propiophenone	20.77	18.49	17.52	16.64	15.11
butyrophenone	51.19	45.71	43.38	41.27	37.61
valerophenone	107.56	96.16	91.32	86.93	79.30
hexanophenone	270.04	241.59	229.49	219.54	199.49
heptanophenone	693.67	619.54	589.73	561.67	512.85
phenylacetamide	2.37	2.02	1.87	1.73	1.49
benzyl alcohol	3.74	3.24	3.03	2.84	2.50
2-phenylethanol	6.31	5.55	5.22	4.92	4.41
3-phenyl-1-propanol	12.67	11.23	10.62	10.07	9.11
4-phenylbutyronitrile	22.22	19.78	18.74	17.81	16.17
toluene	93.85	83.89	79.66	75.82	69.16
methanol-buffer 90:10					
t_o (min)	1.404	1.569	1.652	1.735	1.900
acetophenone	0.75	0.56	0.48	0.41	0.29
propiophenone	0.89	0.69	0.61	0.53	0.40
butyrophenone	1.03	0.82	0.73	0.65	0.50
valerophenone	1.24	1.00	0.90	0.81	0.65
hexanophenone	1.52	1.26	1.15	1.04	0.87
heptanophenone	1.90	1.60	1.47	1.35	1.15
phenylacetamide	0.51	0.35	0.28	0.22	0.12
benzyl alcohol	0.57	0.40	0.33	0.27	0.16
2-phenylethanol	0.61	0.44	0.37	0.31	0.19
3-phenyl-1-propanol	0.67	0.50	0.42	0.35	0.24
methyl 3-phenylpropionate	0.94	0.74	0.65	0.57	0.43
toluene	1.31	1.06	0.96	0.87	0.70

Table 3.9: Effect of changes in the value used for column void volume on calculated capacity factors in acetonitrile eluents

Compound	Calculated capacity factor Column void volume value -15%	-5%	Mean	+5%	+15%
acetonitrile-buffer 30:70					
t_o (min)	0.374	0.418	0.444	0.462	0.506
acetophenone	6.47	5.68	5.29	5.04	4.52
propiophenone	14.85	13.18	12.35	11.83	10.71
butyrophenone	31.28	27.89	26.19	25.13	22.86
valerophenone	68.73	61.39	57.73	55.45	50.54
hexanophenone	155.21	138.77	130.59	125.46	114.46
heptanophenone	351.32	314.24	295.78	284.21	259.41
phenylacetamide	1.19	0.97	0.85	0.77	0.62
benzyl alcohol	2.12	1.79	1.63	1.53	1.31
2-phenylethanol	3.21	2.76	2.54	2.40	2.11
3-phenyl-1-propanol	5.69	4.99	4.64	4.42	3.94
4-phenylbutyronitrile	21.00	18.68	17.53	16.81	15.26
toluene	32.86	29.30	27.52	26.41	24.03
acetonitrile-buffer 90:10					
t_o (min)	1.419	1.586	1.669	1.753	1.919
acetophenone	0.67	0.49	0.42	0.35	0.24
propiophenone	0.80	0.61	0.53	0.45	0.33
butyrophenone	0.94	0.73	0.65	0.57	0.43
valerophenone	1.12	0.90	0.80	0.72	0.57
hexanophenone	1.38	1.13	1.02	0.92	0.76
heptanophenone	1.72	1.40	1.31	1.20	1.01
phenylacetamide	0.49	0.33	0.27	0.21	0.10
benzyl alcohol	0.56	0.40	0.33	0.26	0.15
2-phenylethanol	0.60	0.43	0.36	0.29	0.18
3-phenyl-1-propanol	0.65	0.48	0.41	0.34	0.22
methyl 3-phenylpropionate	0.81	0.55	0.54	0.47	0.34
toluene	1.07	0.85	0.76	0.67	0.53

Table 3.10: Effect of changes in the value used for column void volume on retention indices in methanol eluents

Calculated from capacity factors in Table 3.8.

Compound	Retention index Column void volume value -15%	-5%	Mean	+5%	+15%
methanol-buffer 40:60					
acetophenone	804	804	804	803	803
propiophenone	898	898	898	898	899
butyrophenone	1003	1004	1004	1004	1004
valerophenone	1090	1090	1090	1090	1090
hexanophenone	1197	1197	1197	1197	1197
heptanophenone	1307	1307	1307	1307	1307
phenylacetamide	645	640	638	636	631
benzyl alcohol	698	696	694	693	691
2-phenylethanol	759	758	758	757	756
3-phenyl-1-propanol	840	840	840	840	840
4-phenylbutyronitrile	906	906	906	906	906
toluene	1074	1074	1074	1074	1075
methanol-buffer 90:10					
acetophenone	810	807	804	802	795
propiophenone	906	907	907	908	910
butyrophenone	987	989	991	993	998
valerophenone	1084	1087	1089	1091	1095
hexanophenone	1196	1197	1197	1198	1199
heptanophenone	1317	1313	1311	1308	1303
phenylacetamide	605	578	560	536	455
benzyl alcohol	665	647	635	620	574
2-phenylethanol	705	692	683	673	643
3-phenyl-1-propanol	755	747	741	736	719
methyl 3-phenylpropionate	808	806	803	801	794
toluene	1113	1115	1117	1118	1122

Table 3.11: Effect of changes in the value used for column void volume on retention indices in acetonitrile eluents

Calculated from capacity factors in Table 3.9.

Compound	Retention index Column void volume value -15%	-5%	Mean	+5%	+15%
acetonitrile-buffer 30:70					
acetophenone	799	799	798	798	798
propiophenone	904	904	904	905	905
butyrophenone	998	998	999	999	999
valerophenone	1097	1097	1097	1097	1098
hexanophenone	1199	1199	1199	1199	1199
heptanophenone	1302	1302	1302	1302	1301
phenylacetamide	587	576	569	564	551
benzyl alcohol	659	654	651	649	643
2-phenylethanol	711	708	707	706	703
3-phenyl-1-propanol	783	782	782	782	781
4-phenylbutyronitrile	948	948	948	948	949
toluene	1004	1004	1005	1005	1005
acetonitrile-buffer 90:10					
acetophenone	810	806	803	800	792
propiophenone	903	904	904	905	908
butyrophenone	989	993	995	997	1003
valerophenone	1087	1090	1092	1094	1100
hexanophenone	1196	1197	1197	1198	1198
heptanophenone	1315	1311	1308	1305	1298
phenylacetamide	642	618	602	580	501
benzyl alcohol	714	701	692	680	644
2-phenylethanol	747	738	731	723	700
3-phenyl-1-propanol	796	791	787	783	772
methyl 3-phenylpropionate	913	914	915	917	920
toluene	1061	1064	1066	1068	1074

alkylarylketones (this relationship will be discussed in more detail in Section 3.5).

In the eluents containing 30% acetonitrile and 40% methanol the linearity of the log k' *vs.* C_{no} relationship for the alkylarylketones was not significantly affected by the t_o value with a correlation coefficient ranging from 0.9994 to 0.9995 in methanol and 0.9998 to 0.9999 in acetonitrile. For well retained species falling within the calibrated region of the retention index scale (i.e. $RI > 800$), e.g. toluene, 4-phenyl-1-butyronitrile, the retention indices were virtually independent of the value of t_o used in the calculation, changing by only 1 unit. In these eluents compounds eluting before the calibrated region (i.e. $RI < 800$), benzyl alcohol and phenylacetamide) had retention indices which were more dependent on t_o the effect increasing as the degree of extrapolation increased. However, even with a compound requiring considerable extrapolation (phenylacetamide, RI about 550 in methanol) the retention index changed by only 15 units in methanol and 35 units in acetonitrile.

In the eluents containing 90% organic modifier the correlation coefficients obtained for the relationship between log k' and C_{no} for the alkylarylketones were more sensitive to the value of the column void volume. In methanol the correlation ranged from 0.9977 to 0.9993 and in acetonitrile from 0.9980 to 0.9997. The apparent effect on the linearity in these eluents was also noticeable in the deviations of the retention indices of the alkylarylketones from the nominal values. The well retained species again showed only a small dependence on the t_o value with typical changes in the retention index value of less than 10 units

(toluene in methanol). However, the change in the retention indices of the poorly retained species was in some cases quite large e.g. phenylacetamide 501 to 642 in 90% acetonitrile and increased with the degree of extrapolation. The capacity factors of these compounds were typically less than 0.5 which may suggest that at this level the retention index values may be less robust than at higher retentions although the relative changes were still smaller than for the capacity factors.

Retention indices thus compensate for many of the problems associated with determining the column void volume, if the compound is eluted within the calibrated region, but do not totally overcome the problems when the capacity factors are very small (< 0.5) and the retention index scale extrapolated.

3.3 DEPENDENCE OF CAPACITY FACTORS AND RETENTION INDICES ON THE BUFFER CONCENTRATION OF THE AQUEOUS COMPONENT OF THE MOBILE PHASE

On increasing the organic modifier content of the eluents from 80% to 90% it was necessary to reduce the buffer concentration of the aqueous phase to prevent precipitation of the buffer salt components (sodium dihydrogen orthophosphate and disodium hydrogen orthophosphate). The pH of the buffer was not significantly altered, pH range 6.95 - 7.05, by reducing the concentration. With some model compounds (e.g. aniline, benzyl alcohol, benzyl bromide, benzyl chloride,

nitrobenzene, phenol, phenylacetamide, 4-phenyl-1-butyronitrile, 3-phenyl-1-propanol and 3-phenyl-1-propionitrile) relatively large changes in retention index were observed between 80 and 90% methanol and acetonitrile, as will be discussed in the following chapters. To determine whether the observed changes were due to the enforced change in the buffer concentration and therefore to decide whether the 90% values should be included in the data set, the effect of changing the buffer concentration on the capacity factor and retention index was examined. It was also intended to examine whether the abnormal retention patterns of benzyl amine, 2-phenylethylamine and 3-phenyl-1-propylamine (see Chapter 4) were due to ionic interactions with the stationary phase by examining the effect of the change in ionic strength on the retention of 3-phenyl-1-propylamine in acetonitrile eluents. The full discussion of the retention behaviour of the model compounds will follow in Chapters 4, 5, 6 and 7.

The effects of ionic strength were examined in methanol -buffer (70:30) and acetonitrile-buffer (70:30) eluents in which buffer insolubility should not complicate the observations. The capacity factors of a selection of model compounds and the alkylarylketones were determined using methanol-water, acetonitrile-water, methanol-buffer and acetonitrile-buffer with concentration from 0.001 M to 0.02 M (Table 3.12). The capacity factors were used to calculate retention indices (Table 3.13).

The capacity factors and retention indices of all the compounds, except 3-phenyl-propylamine, appeared to be effectively independent of the buffer concentration. The small changes in observed k' were less than the variations

Table 3.12: Capacity factors obtained using different buffer concentrations as aqueous phases

Compound	Capacity factor Buffer concentration (M) 0.00	0.001	0.002	0.005	0.010	0.020
methanol-buffer 70:30						
benzamide	0.31	0.28	0.30	0.31	0.28	0.27
1,2-dihydroxybenzene	0.32	0.30	0.32	0.32	0.30	0.30
aniline	0.49	0.46	0.47	0.48	0.45	0.43
phenol	0.54	0.51	0.53	0.54	0.50	0.49
m-toluidine	0.73	0.79	0.72	0.73	0.68	0.68
acetophenone[a]	1.00	0.96	0.98	1.01	0.96	0.94
benzene	2.09	2.00	2.04	2.10	2.01	2.03
heptanophenone[a]	11.54	10.98	11.17	11.77	11.31	11.49
acetonitrile-buffer 70:30						
benzamide	0.29	0.30	0.29	0.29	0.29	0.29
1,2-dihydroxybenzene	0.46	0.41	0.41	0.41	0.45	0.45
phenol	0.51	0.51	0.50	0.50	0.50	0.50
aniline	0.59	0.58	0.57	0.58	0.58	0.57
m-toluidine	0.76	0.74	0.74	0.74	0.73	0.74
acetophenone[a]	0.82	0.85	0.86	0.86	0.86	0.85
benzene	1.53	1.49	1.48	1.49	1.51	1.48
heptanophenone[a]	5.20	5.26	5.25	5.37	5.52	5.22
3-phenyl-1-propylamine	[b]	24.23	13.04	8.79	5.52	4.44

[a] the full set of alkylarylketones were examined in the study
[b] not eluted from column.

Table 3.13: Retention indices obtained using different buffer concentrations as aqueous phase

Compound	Retention index Buffer concentration (M) 0.00	0.001	0.002	0.005	0.010	0.020
methanol-buffer 70:30						
benzamide	562	555	558	561	552	546
1,2-dihydroxybenzene	571	568	573	573	567	573
aniline	656	654	653	654	651	645
phenol	677	674	674	676	667	670
m-toluidine	740	738	738	738	733	736
benzene	954	954	954	953	953	955
acetonitrile-buffer 70:30						
benzamide	499	504	501	498	499	494
1,2-dihydroxybenzene	622	590	594	599	621	621
phenol	651	650	650	651	647	647
aniline	690	688	689	691	689	687
m-toluidine	759	756	761	757	755	757
benzene	951	951	953	953	952	952
3-phenyl-1-propylamine	[a]	1733	1558	1440	1320	1259

[a] not eluted from column

Figure 3.1: Change in capacity factors of 3-phenyl-1-propylamine and aniline with aqueous phase buffer concentration

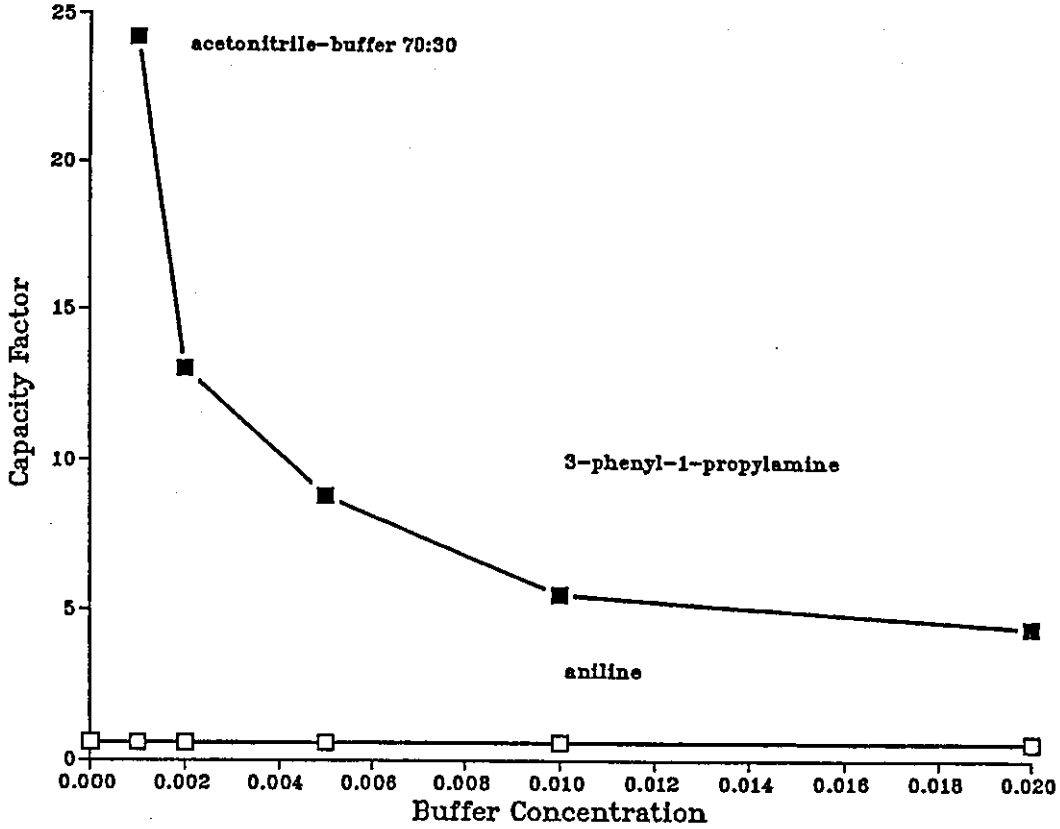
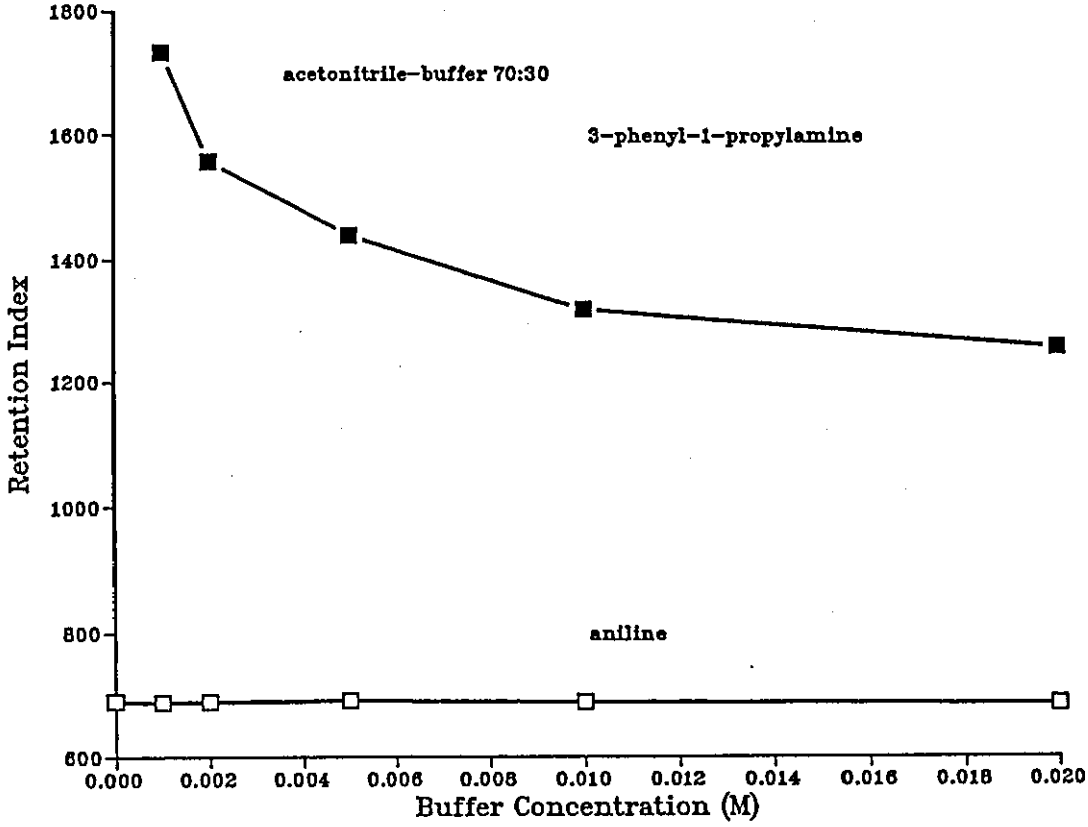


Figure 3.2: Change in retention index of 3-phenyl-1-propylamine and aniline with aqueous phase buffer concentration



observed in a day to day reproducibility study.

3-phenyl-1-propylamine showed a large decrease in both capacity factor (24.33 to 4.44) and retention index (1733 to 1259) with increasing buffer concentration (Figures 3.1 and 3.2). In the eluent containing no buffer this compound was not eluted from the column. This behaviour of 3-phenyl-1-propylamine which has a pKa of 10.39[163] would appear to confirm that in this case the retention mechanism was an ion exchange separation which has previously been suggested for protonated amines[164]. These results contrast with those of aniline, pKa 4.63[165], and m-toluidine, pKa 4.73[165], which would be expected to be un-ionised at this pH and whose retention was not influenced by the buffer concentration.

3.4 REPRODUCIBILITY OF CAPACITY FACTORS

During the study four columns were used and each was packed in the laboratory from the same batch of packing material. To check that the procedures described at the beginning of this chapter were sufficient to produce reproducible capacity factors and to establish the extent of experimental variation the long and short term reproducibility of k' and RI have been examined. The short term reproducibility has been examined using the retention data obtained over a range of eluent compositions for three separate injections on a single day using the same batch of eluent. The long term reproducibility has been studied using two different methods. When each new column was packed the capacity factors and retention indices of a set of test compounds were determined at a single eluent composition

(methanol-buffer 60:40). As part of the main study the retention times of benzene, toluene and phenol were measured on a daily basis and the variations in the observed capacity factors and retention indices calculated.

3.4.1 Short Term Reproducibility of Capacity Factors

The short term reproducibility on a single column has been examined by comparing the results of the three separate determinations on a single day. The capacity factors (Tables 3.14 and 3.15) were determined from a single injection whereas normally they were calculated from the mean of three injections. The study used the same batch of eluent for each eluent composition. The short term reproducibility of the capacity factors was very good,the relative standard deviation from the mean was usually less than 1%.

3.4.2 Long Term Reproducibility of Capacity Factors

The long term reproducibility has been determined using two sets of data, the first of these was for selected compounds which were determined at a single eluent composition on each new column. The second approach was to determine the range, mean and standard deviations from the mean of the k' of benzene, toluene and phenol which were determined on a daily basis.

a) Capacity Factors Measured on Each New Column

The test compounds were chosen to represent the range of strengths of retention on the columns and included poorly

Table 3.14: Reproducibility of capacity factors calculated for a single day in methanol eluents

Mean and standard deviations for three 10 µl injections.

Compound	Methanol proportion (%)					
	40		50		60	
	k'	S.D.	*k'*	S.D.	*k'*	S.D.
acetophenone*	6.12	0.04	2.74	0.01	1.65	0.00
hexanophenone*	195.12	2.02	44.80	0.26	18.08	0.04
benzene	12.13	0.06	5.92	0.02	3.67	0.01
phenol	2.25	0.02	1.14	0.00	0.79	0.01
toluene	29.60	0.16	12.40	0.10	6.54	0.06

	70		80		90	
	k'	S.D.	*k'*	S.D.	*k'*	S.D.
acetophenone*	0.95	0.00	0.58	0.01	0.43	0.01
hexanophenone*	6.29	0.05	2.24	0.00	1.03	0.01
benzene	1.98	0.02	1.08	0.01	0.67	0.00
phenol	0.52	0.01	0.34	0.01	0.24	0.00
toluene	3.33	0.02	1.58	0.01	0.89	0.00

* the full set of alkylarylketones were examined, these are shown as examples

Table 3.15: Reproducibility of capacity factors calculated on a single day in acetonitrile eluents

Mean and standard deviations for three 10 µl injections.

Compound	Acetonitrile proportion (%)							
	30		40		50		60	
	k'	S.D.	*k'*	S.D.	*k'*	S.D.	*k'*	S.D.
acetophenone*	6.04	0.03	3.32	0.01	1.96	0.03	1.19	0.01
hexanophenone*	163.43	0.93	46.87	0.02	16.14	0.30	6.59	0.05
benzene	14.59	0.05	7.66	0.00	3.99	0.02	2.26	0.02
phenol	2.60	0.01	1.60	0.09	1.00	0.03	0.67	0.01
toluene	31.75	0.08	14.23	0.00	6.57	0.01	3.37	0.02

	70		80		90	
	k'	S.D.	*k'*	S.D.	*k'*	S.D.
acetophenone*	0.79	0.01	0.53	0.01	0.33	0.03
hexanophenone*	3.29	0.02	1.55	0.02	0.81	0.02
benzene	1.37	0.01	0.83	0.01	0.47	0.01
phenol	0.45	0.01	0.37	0.01	0.20	0.00
toluene	1.93	0.02	1.08	0.02	0.60	0.01

* the full set of alkylarylketones were examined, these are shown as examples

retained species (phenylacetamide) and well retained species (toluene). The alkylarylketones were also included to enable the calculation of retention indices. Although the columns were packed under the same conditions using the same batch of packing material there were differences in the capacity factors obtained on the four columns (Table 3.16). The relative standards deviations ranged from about 2% to 8% depending on the size of the mean capacity factors, the largest percentage standard deviations from the mean being observed for the longest retained compound (heptanophenone).

b) Capacity Factors of Benzene, Phenol and Toluene

The retentions of benzene, toluene, phenol and the alkylarylketones were measured on a daily basis. At each eluent composition capacity factors were determined for between 20 and 30 individual measurements, on at least two and usually three columns and the mean, range and S.D. calculated (Table 3.17). The results show that there were considerable variations for these compounds across the columns with the maximum and minimum differing by up to 40% from the mean. The variation in capacity factors over the four columns (A to D) is shown clearly for toluene in Figure 3.3 in acetonitrile-buffer 40:60. As with the retention of the selected test compounds described above the largest percentage changes were observed with the strongly retained species (e.g. heptanophenone in 40% methanol).

Table 3.16: Capacity factors of a selection of compounds determined at the same eluent composition on ~~the~~ each of the four columns used in the study

Mobile phase, methanol-buffer(pH7) 60:40.

Compound	Capacity factor Column 1	2	3	4	Mean	S.D.
acetophenone	1.59	1.59	1.54	1.60	1.58	0.02
propiophenone	2.88	2.96	2.74	2.91	2.87	0.08
butyrophenone	4.93	5.18	4.64	4.98	4.93	0.19
valerophenone	8.94	9.62	8.40	9.11	9.02	0.44
hexanophenone	16.61	18.23	15.57	17.01	16.86	0.95
heptanophenone	31.70	34.85	29.38	32.18	32.03	1.94
phenylacetamide	0.52	0.46	0.47	0.47	0.48	0.02
phenol	0.83	0.74	0.74	0.79	0.78	0.04
benzyl alcohol	0.82	0.79	0.78	0.82	0.80	0.02
benzyl cyanide	1.21	1.20	1.15	1.21	1.19	0.02
methyl phenylacetate	2.53	2.30	2.15	2.27	2.31	0.14
benzene	3.57	3.63	3.36	3.53	3.52	0.10
benzyl bromide	5.74	6.01	5.44	5.84	5.76	0.21
toluene	6.52	6.76	6.08	6.57	6.48	0.25

Table 3.17: Reproducibility and range individual measurement of capacity factors of the alkylarylketones and benzene, toluene and phenol in methanol eluents

Compound	Methanol proportion (%)	Capacity factor Mean k'	S.D.	Max. k'	Min. k'
acetophenone*					
	40	6.22	0.47	6.69	4.92
	50	3.02	0.13	3.28	2.87
	60	1.70	0.10	1.84	1.52
	70	0.96	0.04	1.02	0.89
	80	0.63	0.03	0.65	0.57
	90	0.44	0.01	0.46	0.38
hexanophenone*					
	40	189.24	14.10	212.58	155.73
	50	54.49	4.65	60.26	50.30
	60	19.44	2.40	22.49	15.42
	70	6.26	0.34	6.96	5.29
	80	2.47	0.13	2.69	2.30
	90	0.99	0.03	1.04	0.96
benzene					
	40	12.27	0.60	13.22	11.38
	50	6.56	0.50	6.99	5.91
	60	3.55	0.15	3.86	3.22
	70	1.97	0.07	2.09	1.75
	80	1.11	0.04	1.21	1.06
	90	0.65	0.02	0.68	0.61
phenol					
	40	2.24	0.21	2.80	2.02
	50	1.26	0.10	1.29	1.14
	60	0.71	0.02	0.83	0.75
	70	0.59	0.02	0.57	0.50
	80	0.34	0.02	0.39	0.33
	90	0.25	0.01	0.27	0.25
toluene					
	40	29.83	1.07	31.80	27.25
	50	13.92	0.90	16.07	12.32
	60	7.13	0.84	7.81	5.54
	70	3.24	0.28	3.46	2.40
	80	1.66	0.07	1.80	1.56
	90	0.86	0.03	0.89	0.81

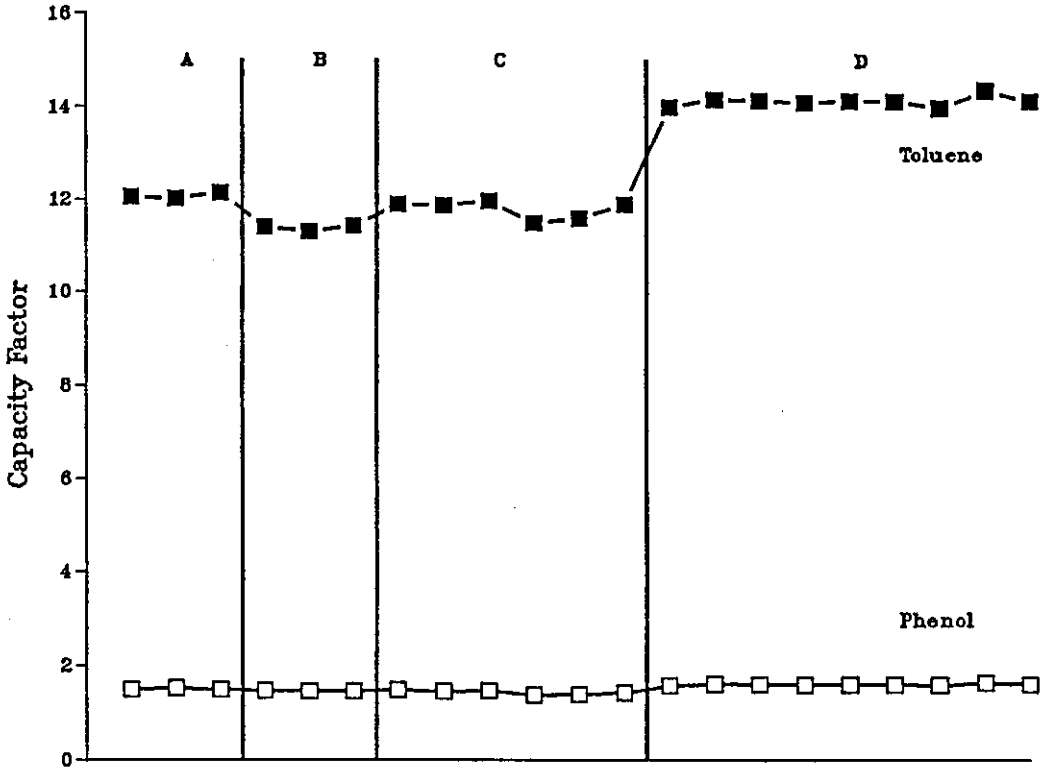
* the complete set of alkylarylketones were measured, these are shown as examples

Table 3.18: Reproducibility and range of individual measurement of capacity factors of the alkylarylketones and a limited group of test compounds in acetonitrile eluents

Compound	Acetonitrile proportion (%)	Capacity factor Mean k'	S.D.	Max. k'	Min. k'
acetophenone*					
	30	5.81	0.41	6.71	5.24
	40	3.07	0.20	3.36	2.79
	50	1.81	0.12	2.08	1.67
	60	1.15	0.04	1.21	1.09
	70	0.77	0.03	0.82	0.72
	80	0.54	0.05	0.61	0.46
	90	0.33	0.02	0.36	0.30
hexanophenone*					
	30	144.61	15.05	167.84	123.03
	40	40.60	5.25	47.58	36.04
	50	14.03	1.50	16.52	11.55
	60	6.02	0.48	6.29	5.53
	70	3.07	0.21	3.33	2.79
	80	1.67	0.13	1.88	1.51
	90	0.79	0.03	0.83	0.73
benzene					
	30	13.76	1.02	15.38	11.99
	40	6.90	0.57	7.79	6.14
	50	3.70	0.37	4.27	2.98
	60	2.14	0.10	2.28	2.03
	70	1.31	0.05	1.40	1.29
	80	0.84	0.06	0.95	0.74
	90	0.48	0.02	0.51	0.44
phenol					
	30	2.46	0.21	2.84	1.86
	40	1.51	0.07	1.64	1.38
	50	0.96	0.18	1.09	0.84
	60	0.65	0.02	0.68	0.63
	70	0.44	0.01	0.47	0.42
	80	0.31	0.05	0.39	0.24
	90	0.21	0.01	0.21	0.20
toluene					
	30	29.75	2.19	31.98	25.27
	40	12.78	1.20	14.37	11.31
	50	6.12	0.59	7.01	4.94
	60	3.18	0.15	3.41	3.01
	70	1.87	0.08	1.98	1.73
	80	1.13	0.09	1.26	1.09
	90	0.60	0.03	0.64	0.56

* the complete set of alkylarylketones were measured, these are shown as examples

Figure 3.3: Variation in capacity factor of toluene and phenol on the four columns used in the study



Despite the precautions noted earlier to ensure that conditions remained as constant as possible the variation of the capacity factors in this long term study was considerably larger than the variation on a single day. The results show that this "absolute" method of reporting retention needs to be used with care when the study is a long term, one in particular where more than one column has to be used.

3.5 CALCULATION OF RETENTION INDICES

The study discussed in the previous sections showed that the use of capacity factors for the long term collection of data may present problems. This was particularly significant when more than one column was used to collect the retention data. Previous work using retention indices has shown that they were considerably more robust to the column void volume value, the exact experimental conditions and small changes in eluent composition than capacity factors as they provided an interpolated scale in which the retention of a compound was expressed relative to the standard alkylarylketones. In the following sections the linearity of the assumed log k' vs. C_{no} relationship which forms the backbone of the calculation of retention indices will be investigated and the reproducibility of retention indices will also be studied.

Retention indices were calculated as described by Smith[134] from the least squares regression equation obtained for log k' vs. 100 x C_{no} for the alkylarylketone standards (acetophenone - heptanophenone, C_8 - C_{13}). The retention

indices on a particular day were calculated using the capacity factors for the alkylarylketones included in that set of runs of the test compounds. The calculations below were for a typical set of capacity factors (Table 3.19), these were not related to the values used previously in the calculation of the column void volume (Section 3.2). The differences again emphasise the problems of using k' to report retention.

3.5.1 Linearity of log Capacity Factor - Carbon Number Relationship for Alkylarylketones

The correlations for the regression equations relating log k' and 100 x C_{no} were high for all the eluent compositions and showed a good linearity (Table 3.20, Figures 3.4 and 3.5). Correlation coefficients between 0.9994 and 0.9999 were observed for the eluents containing 30 to 80 % organic modifier. There was a reduction in the correlation coefficient for the eluent containing 90% methanol to 0.9985 suggesting the possibility of some curvature at this eluent composition. The capacity factors at this eluent composition have been shown to be sensitive to the t_o value used in their calculation. The possibility of curvature was also shown in the systematic deviations of the calculated retention index values for the alkylarylketones standards from their nominal values (Table 3.21).

The slope of the log k' vs. C_{no} relationship is a measure of the methylene group selectivity of the system. Theoretically the slope should be the same for different homologous series although the intercept will vary depending

on the functional groups present. However, some workers have suggested that the methylene group selectivity differs between different homologous series[4] due to the influence of the functional group.

It has been suggested that the lines for log *k'* *vs.* carbon number at different eluent compositions should converge to a single point in methanol and 2 points in acetonitrile[162]. The single convergence point in methanol was seen as an indication of a linear relationship between the slope and methanol concentration. The dual convergence points in acetonitrile have been interpreted as representing a two stage curve with a discontinuity at 50% acetonitrile. Previous work has found that the x axis co-ordinate of the convergence point, that is the carbon number, was dependent on the homologous series but that the y co-ordinate (log *k'*) was independent[162]. The possibility of a convergence point has been examined by extrapolating the data for log *k'* *vs.* 100 x C_{no} for the alkylarylketones to an equivalent of 0 carbon atoms. The results (Figures 3.6 and 3.7) showed that in methanol all the lines except 40% methanol appear to converge to a single point at a carbon number equivalent of approximately 4.2. In acetonitrile all the curves except the 90% one converge to a single point at 3.4-3.6 carbon numbers although 30 and 40% lines may converge to a slightly higher value. From this data it was not possible to identify two definite convergence points.

3.5.2 Relationship Between Slope and Intercept of Correlation Curves and Eluent Composition

The relationship between slope and modifier

concentration, and intercept and modifier concentration is a useful measure which will be used to back-calculate capacity factors from the calculated retention indices in the prediction system (Chapter 8). The mean slope and intercept calculated for the 20 - 30 individual determinations at each eluent composition over the period of the study (Tables 3.17 and 3.18) were determined (Table 3.22). The values used therefore differ from those in Table 3.20 which were based on the values determined on a single column in a single set of runs (at each eluent composition). The relationship between mean slope and methanol concentration was approximately linear, however the equivalent for acetonitrile showed definite curvature (Figure 3.8). Jandera and co-workers[120,166] have also found the relationship between slope and methanol concentration was linear with alkylbenzenes but curved for n-alkanes. In acetonitrile the slope - eluent composition curves were found to be non-linear for both homologous series. In the present study the curves for intercept and eluent composition were also found to be approximately linear for methanol eluents but curved for acetonitrile (Figure 3.9).

The change in slope with eluent composition and the change in intercept with eluent composition can be described by quadratic regression equations (Table 3.23). In this Table the coefficients have been calculated taking the mean slope and intercept values from Table 3.22.

Figure 3.4: Linearity of the log capacity factor - carbon number relationship for the alkylarylketone standards in methanol eluents

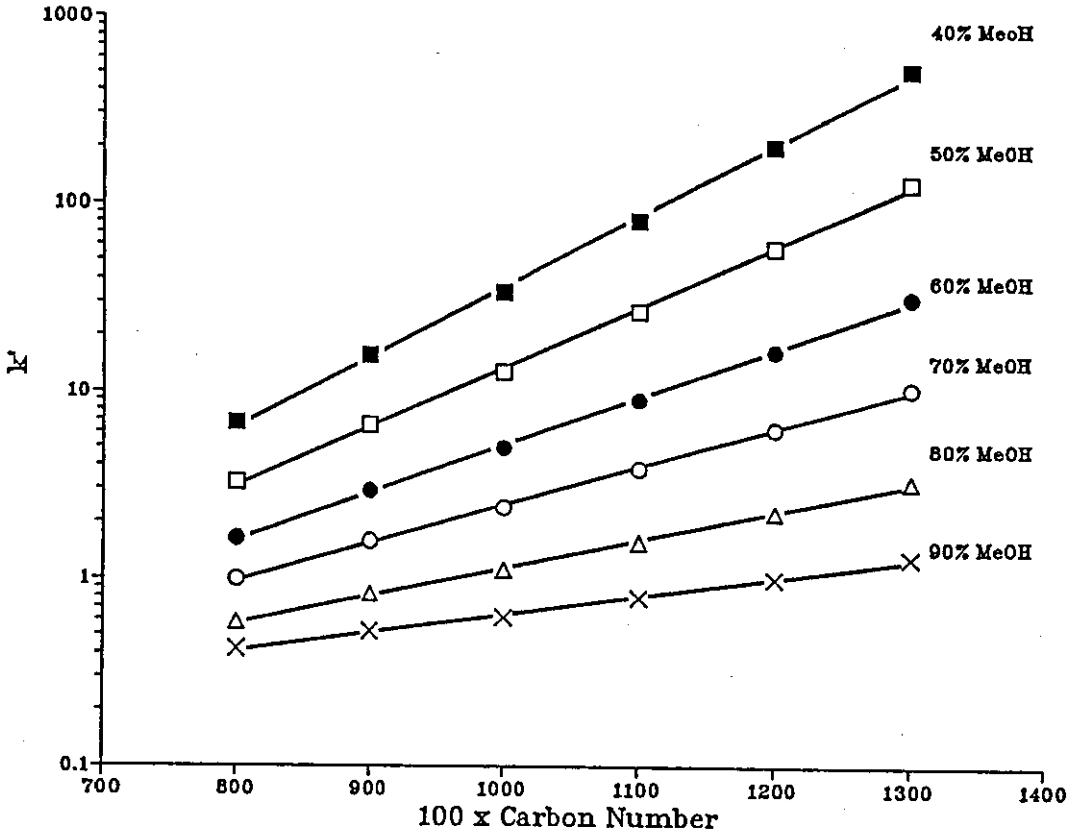


Figure 3.5: Linearity of the log capacity factor - carbon number relationship for the alkylarylketone standards in acetonitrile eluents

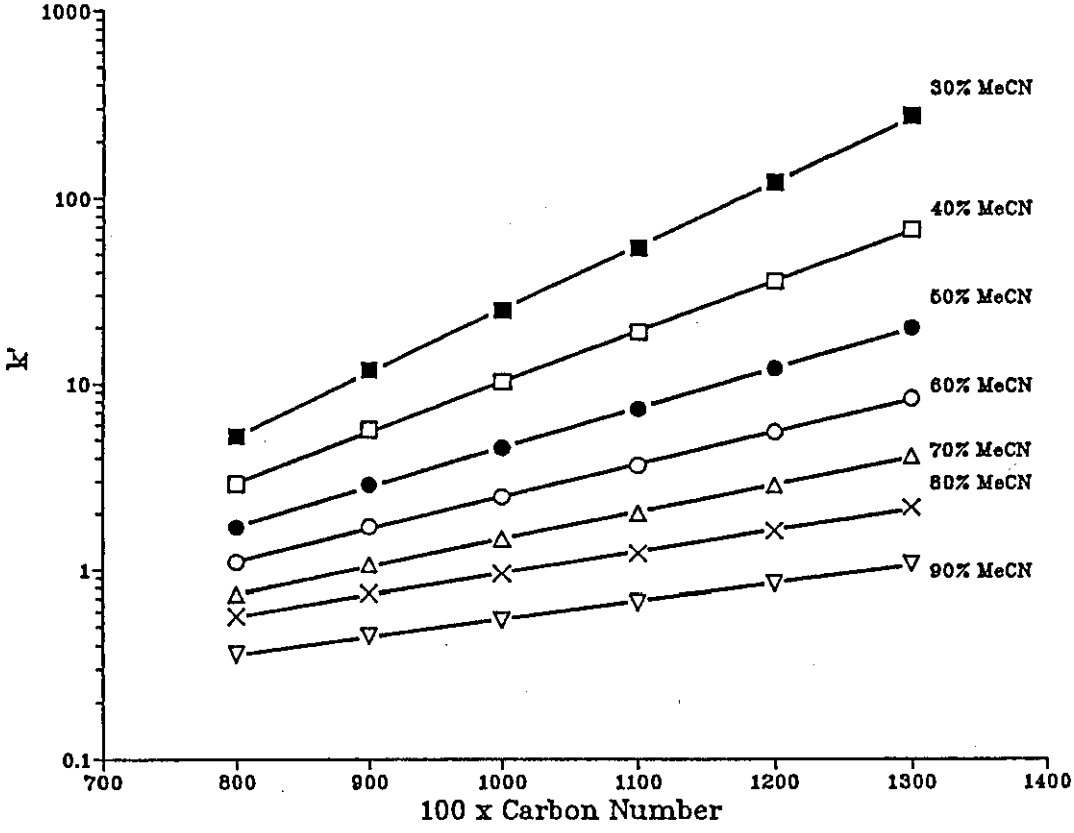


Table 3.19: Capacity factors of a typical set of alkylarylketone standards on a single column

Compound	Capacity factor (k') Modifier proportion (%) 30	40	50	60	70	80	90
methanol-buffer							
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
acetonitrile-buffer							
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

Table 3.20 : Coefficients and correlation of regression equations for log k'- 100 x carbon number for alkylarylketones

Based on capacity factors in Table 3.19.

$$\log k' = a\ (100 \times C_{no}) + b$$

Modifier proportion (%)	Coefficients of regression equations a (x 10^{3})	b	Correlation coefficient
methanol-buffer			
40	3.778	-2.132	0.9994
50	3.214	-2.080	0.9995
60	2.581	-1.865	0.9995
70	2.053	-1.658	0.9995
80	1.492	-1.434	0.9994
90	0.985	-1.179	0.9985
acetonitrile-buffer			
30	3.441	-2.033	0.9999
40	2.734	-1.718	0.9999
50	2.143	-1.481	0.9998
60	1.757	-1.359	0.9998
70	1.478	-1.308	0.9997
80	1.164	-1.179	0.9994
90	0.955	-1.213	0.9994

Table 3.21: Calculated retention indices for the alkylarylketone standards

Based on capacity factors in Table 3.19.

Compound	*RI**	Retention index Modifier proportion (%) 30	40	50	60	70	80	90
methanol-buffer								
acetophenone	800	-	806	805	805	805	802	810
propiophenone	900	-	902	902	903	904	907	905
butyrophenone	1000	-	993	993	993	993	994	983
valerophenone	1100	-	1093	1094	1095	1093	1092	1091
hexanophenone	1200	-	1198	1199	1197	1199	1198	1197
heptanophenone	1300	-	1308	1307	1308	1307	1308	1313
acetonitrile-buffer								
acetophenone	800	800	798	797	797	797	803	806
propiophenone	900	904	905	906	906	905	906	907
butyrophenone	1000	997	999	999	999	998	998	998
valerophenone	1100	1097	1098	1097	1096	1094	1093	1095
hexanophenone	1200	1199	1199	1199	1199	1198	1197	1196
heptanophenone	1300	1303	1302	1303	1304	1305	1308	1309

* nominal retention index values

Figure 3.6: Extrapolation of log capacity factor vs. carbon number relationship to show convergence point in methanol eluents

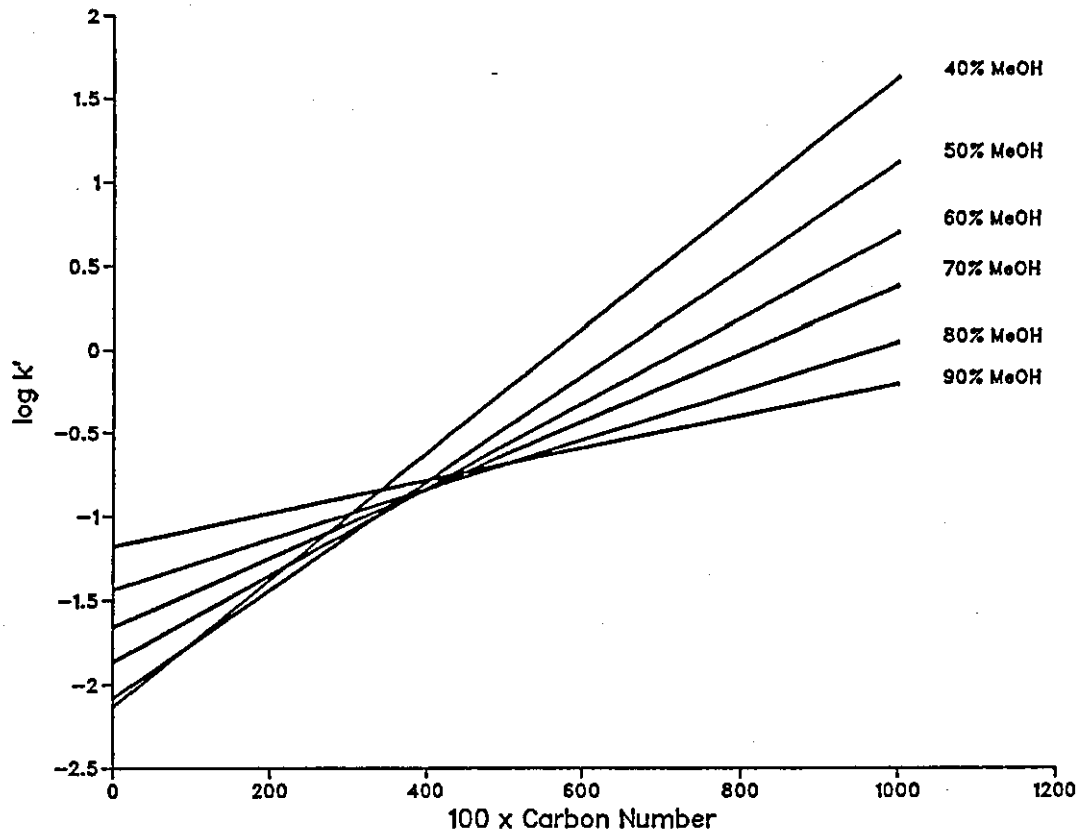


Figure 3.7: Extrapolation of log capacity factor vs. carbon number relationship to show convergence point in acetonitrile eluents

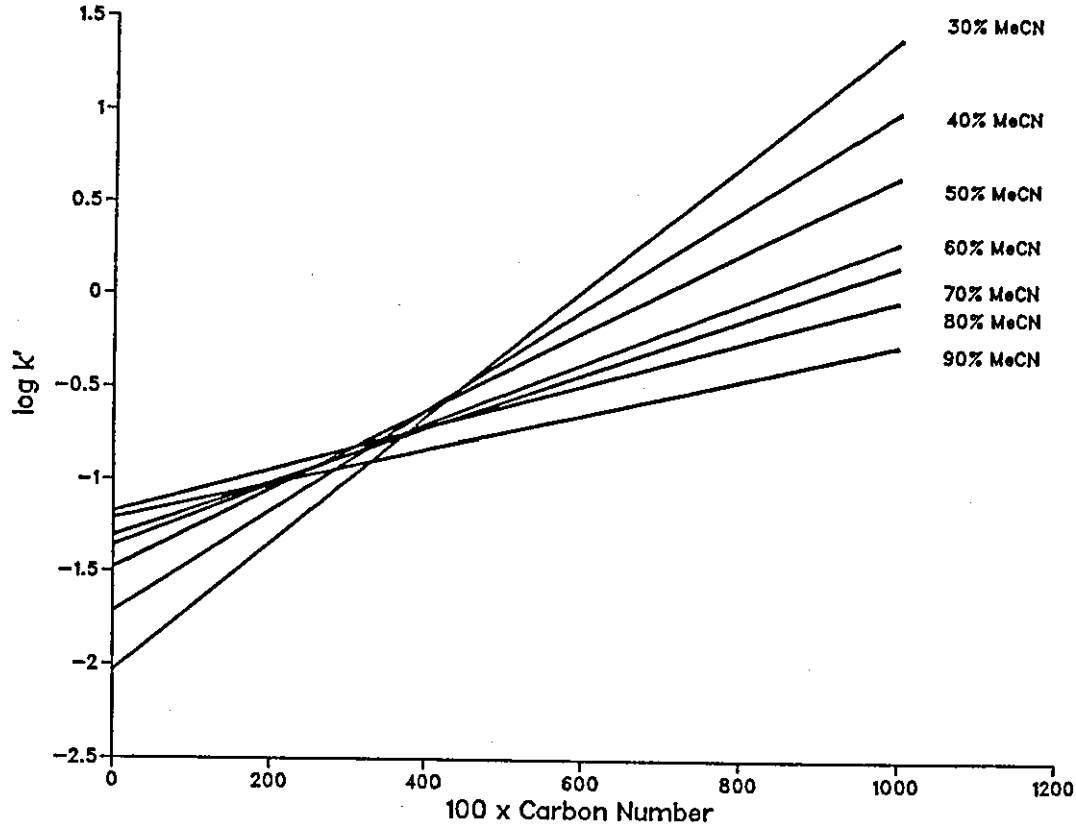


Table 3.22: Reproducibility and linearity of the log *k'* vs. carbon number relationship for the alkylarylketones

$$\log k' = a\ (100 \times C_{no}) + b$$

Modifier-buffer	Coefficients of regression equations			
	Slope (a)		Intercept (b)	
	Mean x 10^{3}	S.D. x 10^{3}	Mean	S.D.
methanol-buffer				
40:60	3.762	0.04	-2.239	0.05
50:50	3.160	0.08	-2.074	0.04
60:40	2.665	0.09	-1.914	0.04
70:30	2.045	0.06	-1.657	0.04
80:20	1.514	0.06	-1.434	0.06
90:10	0.964	0.07	-1.154	0.08
acetonitrile-buffer				
30:70	3.500	0.05	-2.045	0.02
40:60	2.776	0.06	-1.729	0.03
50:50	2.184	0.07	-1.484	0.03
60:40	1.778	0.04	-1.361	0.02
70:30	1.482	0.04	-1.298	0.04
80:20	1.235	0.05	-1.248	0.07
90:10	0.960	0.02	-1.234	0.04

Table 3.23: Relationship between slope and intercept of log *k'* vs. carbon number of alkylarylketones (Table 3.32) and organic modifier concentration

$$y = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Modifier	Coefficients of quadratic regression equations					
	Slope			Intercept		
	a (x 10^{6})	b (x 10^{5})	c (x 10^{3})	a (x 10^{4})	b (x10^{2})	c
methanol	0.136	-5.776	6.043	1.521	0.201	-2.563
acetonitrile	0.637	-11.471	6.354	3.675	5.586	-3.381

Figure 3.8: Relationship between mean slope and organic modifier proportion

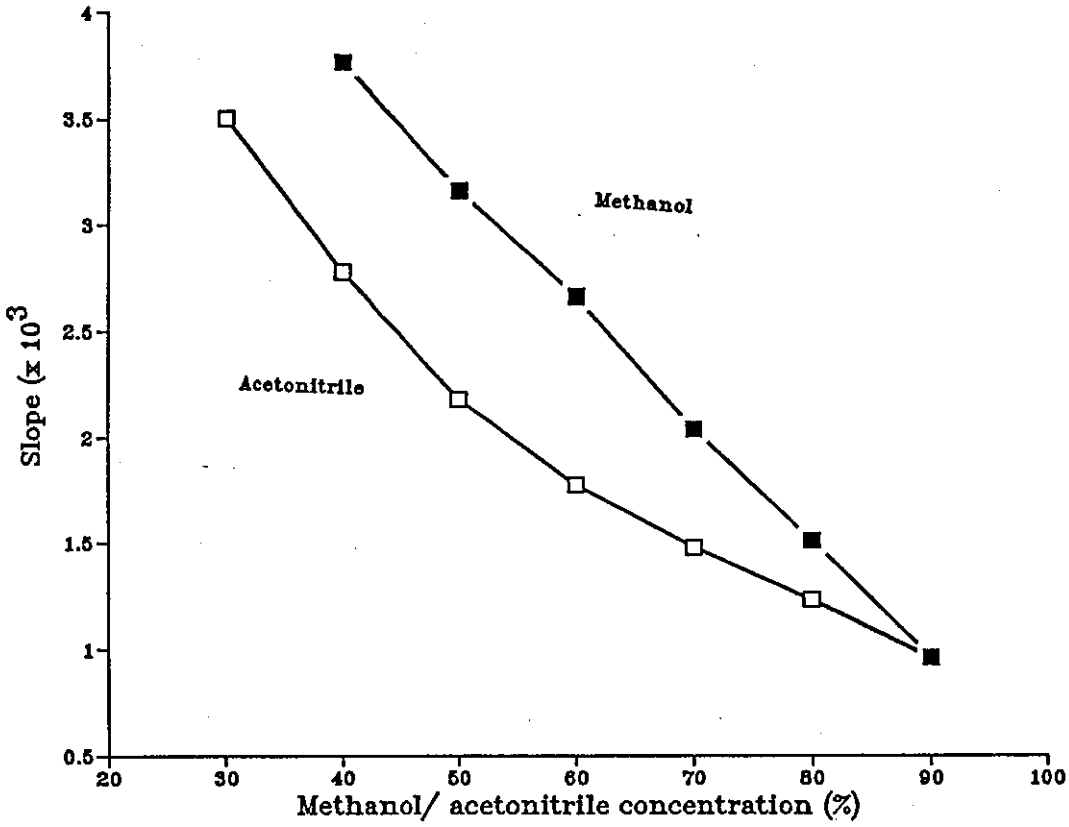
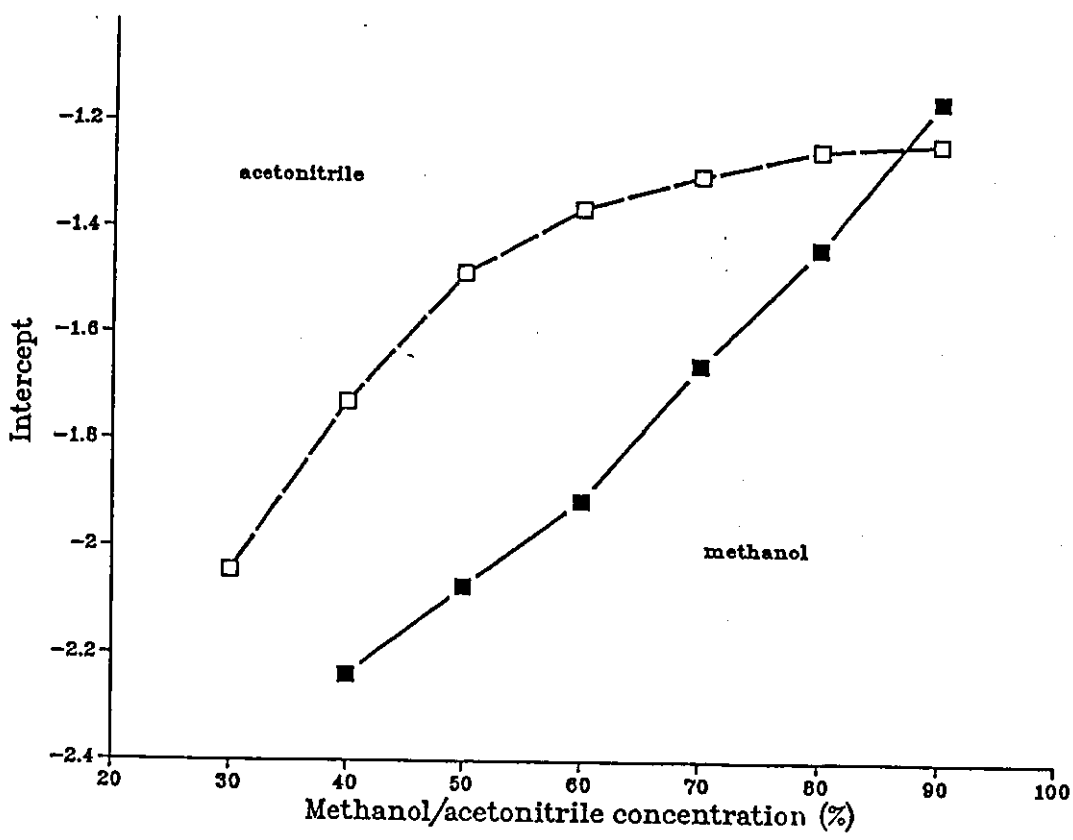


Figure 3.9: Relationship between mean intercept and organic modifier proportion



3.6 REPRODUCIBILITY OF RETENTION INDICES

Using the data reported previously (Section 3.4) the retention indices of the test compounds were calculated from the linear relationship between log k' and 100 x C_{no} for the alkylarylketones. The reproducibility of the retention indices on a long and short term basis has been examined as described for the capacity factors in Section 3.4. The intention was to examine whether retention indices could be used to provide a more reproducible method of reporting retention than the capacity factors.

3.6.1 Short Term Reproducibility of Retention Indices

The short term reproducibility of the retention indices was very good (Table 3.24 and 3.25). As with the capacity factors these retention indices were calculated on the basis of single injections rather than the mean of three injections and would therefore be expected to represent the worst possible reproducibility on a single day. The observed standard deviations from the mean were higher for the 90% methanol and acetonitrile values for which the absolute capacity factors were small and would therefore be very dependent minor measurement errors. With the exception of the 90% values the relative standard deviations are less than 1% showing the reproducibility of the method and suggest an expected uncertainty in a single retention index value of less than ± 4 units.

Table 3.24: Reproducibility of retention indices determined on a single day for methanol eluents

Compound	Retention index Methanol proportion (%) 40		50		60		70		80		90	
	RI	S.D.	*RI*	S.D.	*RI*	S.D.	*RI*	S.D.	*RI*	S.D.	*RI*	S.D.
acetophenone	805	0.5	805	1.2	804	0.0	804	0.0	804	0.0	805	0.0
propiophenone	902	0.5	902	0.5	903	0.0	904	0.0	905	0.0	907	0.8
butyrophenone	993	0.0	993	0.8	993	0.0	993	0.0	992	0.5	990	0.5
valerophenone	1094	0.0	1095	0.7	1094	0.0	1094	0.5	1092	0.0	1090	0.5
hexanophenone	1199	0.0	1201	0.8	1199	0.0	1199	0.0	1198	0.0	1197	0.0
heptanophenone	1303	0.1	1307	0.0	1307	0.0	1307	0.0	1309	0.5	1311	0.5
benzene	883	0.5	915	1.4	935	0.5	955	0.5	983	2.6	999	3.7
phenol	692	0.5	681	3.3	684	2.5	671	1.3	651	5.7	589	5.2
toluene	984	0.5	1009	2.2	1038	0.8	1062	0.5	1094	0.9	1132	4.5

Table 3.25: Reproducibility of retention indices determined on a single day for acetonitrile eluents

Compound	Retention Index Acetonitrile proportion (%) 30		40		50		60	
	RI	S.D.	*RI*	S.D.	*RI*	S.D.	*RI*	S.D.
acetophenone	798	0.0	797	0.0	796	0.0	796	0.0
propiophenone	904	0.0	905	0.0	906	0.0	907	0.5
butyrophenone	999	0.0	1000	0.0	1000	0.5	1000	0.0
valerophenone	1098	0.0	1098	0.0	1098	0.5	1097	0.0
hexanophenone	1201	0.0	1199	0.0	1199	0.0	1198	0.0
heptanophenone	1301	0.1	1301	0.0	1300	0.5	1302	0.0
benzene	908	3.9	926	4.4	937	2.9	949	2.2
phenol	695	2.2	688	3.2	672	4.4	660	1.2
toluene	1003	2.9	1019	3.8	1031	4.4	1043	2.6

	70		80		90	
	RI	S.D.	*RI*	S.D.	*RI*	S.D.
acetophenone	797	0.0	803	0.5	801	5.9
propiophenone	907	0.0	905	0.8	905	6.0
butyrophenone	1000	0.5	995	0.5	997	0.5
valerophenone	1096	0.0	1092	0.8	1092	0.9
hexanophenone	1198	0.0	1197	0.0	1197	0.5
heptanophenone	1303	0.0	1308	0.0	1307	1.6
benzene	955	3.5	962	3.8	971	6.5
phenol	641	2.1	618	1.8	589	4.9
toluene	1051	3.5	1061	3.2	1077	2.9

3.6.2 Long Term Reproducibility

The data reported for the long-term reproducibility of capacity factors (Section 3.4.2) was used to calculate retention indices and therefore to examine the reproducibility over the period of the study.

The retention indices for the set of compounds used to test the four columns are given in Table 3.26 and for benzene, toluene and phenol in Tables 3.27 and 3.28. The results suggest that the retention indices can be expressed as the measured value ± 10 units (twice the standard deviation, the 90% confidence limits). This would account for most of the observed variations in retention index. Where the capacity factors are very small (80 and 90% MeOH and MeCN) and in particular where the retention index scale was extrapolated, i.e. phenol, a slightly larger uncertainty in the value might be expected. The size of the uncertainty in any retention index can be used to determine whether differences between predicted and experimental retention indices were significant.

The results suggest that retention indices were considerably more reproducible than capacity factors even when the experimental parameters were carefully controlled. However, a direct comparison of the relative standards deviations of the two scales would not be appropriate as the scales involved were considerably different. The retention indices were more robust to the uncertainty in the column void volume and differences in the overall retention of the column.

Table 3.26: Reproducibility of retention indices of a selection of compounds determined at a single eluent composition on the four columns used in the study

Mobile phase, methanol-buffer 60:40.

Compound	Retention index Column 1	2	3	4	Mean	S.D.
phenylacetamide	605	600	604	598	602	2.9
phenol	680	678	680	685	681	2.6
benzyl alcohol	694	689	690	692	691	1.9
benzyl cyanide	758	756	756	757	757	0.8
methyl phenylacetate	861	863	862	862	862	0.7
benzene	904	937	938	936	938	1.5
toluene	1041	1038	1039	1040	1040	1.1
benzyl bromide	1020	1019	1020	1020	1020	0.4

Table 3.27: Reproducibility of retention indices of the benzene toluene and phenol determined on four columns over the period of the study, in methanol eluents

Compound	Methanol proportion (%)	Retention index Mean	Max.	Min.	S.D.
benzene					
	40	885	891	883	5.6
	50	913	917	909	4.0
	60	936	941	933	2.0
	70	956	962	952	2.0
	80	980	986	977	2.8
	90	1001	1004	995	5.4
phenol					
	40	691	704	685	5.4
	50	689	694	681	7.0
	60	683	689	679	2.7
	70	673	679	669	3.2
	80	654	663	648	5.4
	90	586	594	561	7.1
toluene					
	40	986	989	979	5.7
	50	1015	1022	1012	3.7
	60	1036	1045	1031	4.3
	70	1063	1065	1059	1.8
	80	1091	1097	1088	2.9
	90	1128	1136	1122	7.6

Table 3.28: Reproducibility of retention indices of benzene, toluene and phenol determined on the four columns used for the study in acetonitrile eluents

Compound	Acetonitrile proportion (%)	Retention index Mean	Max.	Min.	S.D.
benzene					
	30	905	910	899	3.9
	40	926	932	923	4.4
	50	939	941	931	2.9
	60	949	952	944	2.2
	70	956	959	951	3.5
	80	962	969	958	3.8
	90	971	979	964	6.5
phenol					
	30	695	697	688	2.2
	40	688	696	682	3.2
	50	672	682	664	4.4
	60	660	661	658	1.2
	70	641	642	638	2.1
	80	625	652	611	11.6
	90	589	595	582	4.9
toluene					
	30	1003	1008	999	2.9
	40	1019	1026	1017	3.8
	50	1031	1037	1024	4.4
	60	1043	1047	1039	2.6
	70	1051	1057	1049	3.5
	80	1061	1065	1055	3.2
	90	1077	1083	1074	2.9

CHAPTER 4

PARENT INDEX AND SUBSTITUENT INDICES FOR SINGLE SUBSTITUENTS ON ALIPHATIC SIDE CHAINS

In the retention prediction method described in this work the intention was to calculate retention indices from molecular structure as the sum of the parent contribution (PI), the contribution of the aliphatic substituents (SI_{Al-x}), the contribution of the aromatic substituents (SI_{Ar-x}), the contribution of an alkyl chain (SI_{R}) and terms to account for any interactions between substituents (II_{y-z}), according to the equation:-

$$RI = PI + SI_{R} + \Sigma SI_{Al-x} + \Sigma SI_{Ar-x} + \Sigma II_{y-z} \qquad 4.1$$

Each term in this equation has been described using a quadratic equation to describe the change with eluent composition. The following four chapters are concerned with collecting the data which will be used as the basis of the retention prediction method.

Throughout the study a single parent compound, benzene, has been used. The substituted derivatives of this compound could be readily determined by UV absorption and a large number of derivatives were readily available. The use of the UV detector prevented the substituent indices being obtained for purely aliphatic compounds. Substituted alkylbenzenes with substituents as terminal substituents on the alkyl chain have therefore been used to obtain the SI values for aliphatic substituents. In the absence of any long range interactions with the aromatic ring it would be expected

that these values should be similar to those obtained from substituents on purely aliphatic compounds but this has not been studied. The substituent indices have also been determined for a range of model compounds with single substituents an aromatic ring (Chapter 6).

The retentions of all these compounds were measured over the eluent compositions methanol-buffer 40:60 to 90:10 v/v and acetonitrile-buffer 30:70 - 90:10 v/v. In some of the later studies the retentions were not measured with 90% organic modifier due to the problems of uncertainty of the determined values (discussed later in this Chapter).

In this chapter the determination of the parent index values and the substituent index values for aliphatic substituents will be described. Whenever possible substituents have been examined as terminal substituents on more than one length of alkyl chain so that any interactions with the ring could be investigated. The three alkylbenzenes on which the functional groups were substituted were toluene, ethylbenzene and n-propylbenzene. In all subsequent studies the standardised procedure described in the previous chapter was adopted.

4.1 DEFINITION AND CALCULATION OF PARENT INDEX VALUES

Throughout the study benzene has been used as the parent compound. The parent index value in the above equation (4.1) was defined as the value calculated from a quadratic equation relating the change in retention index to % organic modifier of the mobile phase.

The capacity factors of benzene over the eluent ranges

methanol (40-90%) and acetonitrile (30-90%) were therefore measured (Table 4.1). The capacity factors were used to calculate retention indices (Table 4.2) using the linear relationship between log k' and 100 x carbon number of the alkylarylketone standards (Section 3.5.1). Rather than use the individual empirical retention index values as the basis of the study it was decided to fit the points to a quadratic equation and use the calculated, and therefore smoothed, values as the basis of the study (Table 4.3). The change in the retention index in each eluent range (up to 80% organic modifier) was found to be well described using quadratic regression equations (Table 4.2, Figure 4.1). As described later in this chapter it was decided to restrict the study to this range of modifier. These regression equations were important in the prediction scheme as they provide the reference values for the parent retention index which can then be used to calculate the substituent index.

4.2 RETENTION BEHAVIOUR OF ALIPHATIC SUBSTITUTED ALKYLBENZENES

The retention of three alkylbenzenes, toluene, ethylbenzene and n-propylbenzene was measured for two purposes, firstly to determine whether the methylene increment, which in the alkylarylketones is defined as 100, was also valid for the alkylbenzene homologous series. Secondly the alkylbenzenes would be the parent compound for the aliphatic substituted model compounds. By comparing the three sets of model compounds with different chain lengths

Table 4.1: Capacity factors for benzene in different eluents

Typical sets of data determined on a single column

Modifier	Capacity factor Organic modifier proportion (%)						
	30	40	50	60	70	80	90
methanol	-	12.33	7.38	3.58	2.01	1.08	0.64
acetonitrile	14.17	6.53	3.87	2.26	1.33	0.93	0.45

Table 4.2: Retention indices of benzene determined in different eluents

Modifier	Retention index Organic modifier proportion (%)						
	30	40	50	60	70	80	90
Empirical values							
methanol	-	883	915	938	958	983	999
acetonitrile	910	927	940	951	960	963	962
Parent index values							
methanol*	-	885	913	938	961	982	-
acetonitrile*	910	927	940	951	958	963	-

* Parent index values calculated from regression equations in Table 4.3

Table 4.3: Regression equations for the change in retention index of benzene with eluent composition.

These equations were used to calculate *PI*

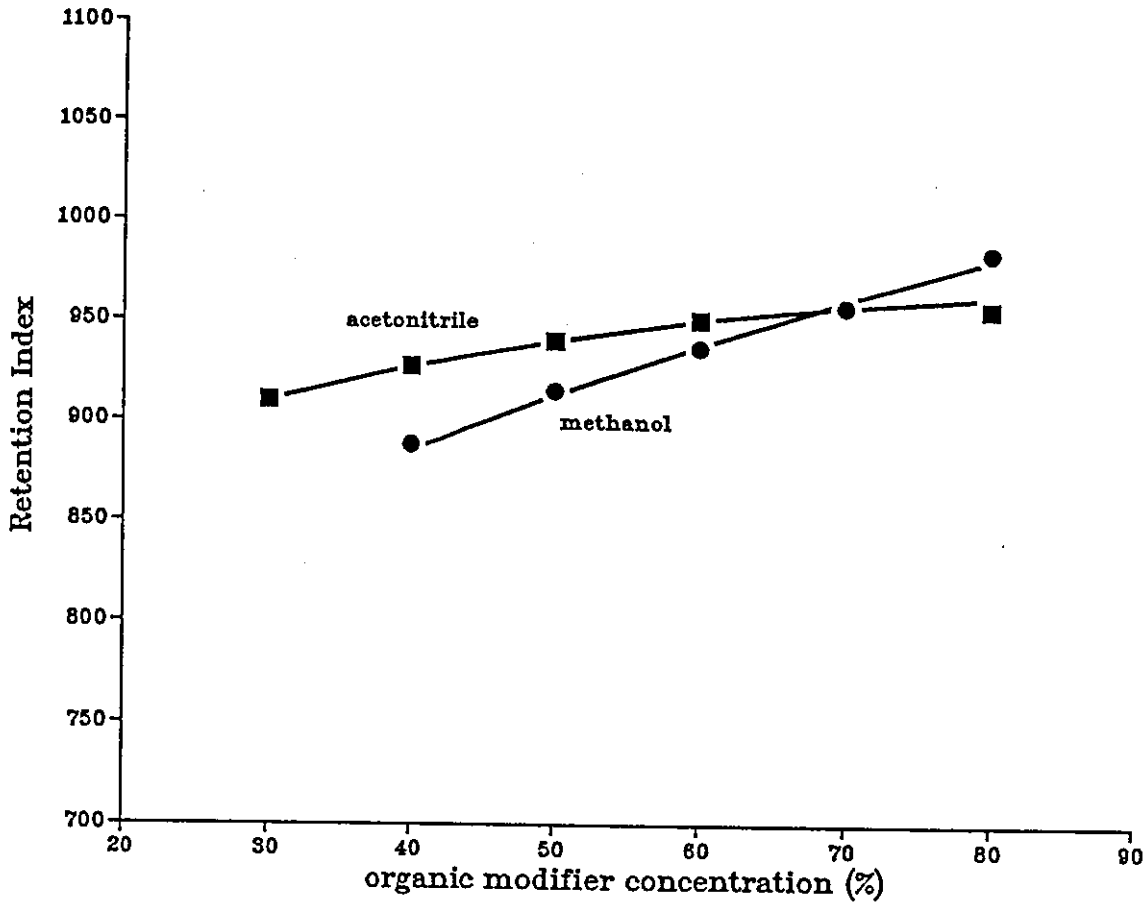
$$PI = ax^2 + bx + c$$

x = % modifier

Modifier	Range (%)	Coefficients of *PI* equations a	b	c
MeOH	40 - 80	-0.0121	3.887	748
MeCN	30 - 80	-0.0154	2.761	841

Figure 4.1: Comparison of experimental retention indices of benzene and calculated parent indices

The points are the experimental values (Table 4.2) and the lines are the values calculated from the regression equations for parent index in Table 4.3.



any interactions between the substituent and the aromatic ring could be observed. The retentions of the substituents hydroxyl, bromide, chloride, nitrile, and methyl carboxylate were measured on all three homologues. However, the amide, aldehyde and ether functional groups could only be measured on one or two of the alkylbenzenes.

4.2.1 Capacity Factors of Model Compounds

The capacity factors (Table 4.4 and 4.5) of the alkylbenzenes and a wide range of substituted model compounds were determined. These values may have been obtained on different days and or on different columns and close and direct comparisons are not therefore possible. With the exception of the three amines, benzylamine, 2-phenylethylamine and 3-phenyl-1-propylamine, capacity factors decreased systematically with increasing organic modifier concentration. The unusual behaviour of these amines will be discussed in Section 4.3.3. The change in capacity factor with eluent composition was examined for a selection of the compounds (Figures 4.2 and 4.3). For some compounds (e.g. 3-phenyl-1-propyl bromide), the curve between log k' and % methanol was approximately linear (Figure 4.2) in agreement with the findings of Schoenmakers[59] but in acetonitrile (Figure 4.3) all the curves were definitely non-linear. Smith *et al.*[145] found similar relationships in acetonitrile however Snyder and co-workers report the successful use of linear relationships in acetonitrile[167,168,169] for retention prediction.

Table 4.4: Capacity factors for compounds with an aliphatic substituent on the alkyl chain in eluents containing methanol

Compound	Capacity factor (*k'*) Methanol proportion (%)					
	40	50	60	70	80	90
toluene	29.51	14.02	7.37	3.37	1.74	0.84
benzyl alcohol	2.31	1.31	0.84	0.56	0.41	0.23
benzylamine	4.64	2.42	1.50	0.87	0.60	0.36
benzyl bromide	31.54	13.40	6.52	2.85	1.56	0.49
benzyl chloride	24.56	10.90	5.51	2.38	1.24	0.49
benzyl cyanide	4.79	2.28	1.31	0.73	0.48	0.26
methyl phenylacetate	10.39	4.68	2.53	1.25	0.74	0.40
phenylacetaldehyde	5.35	2.01	1.13	0.70	0.56	--
phenylacetamide	1.36	0.79	0.52	0.38	0.30	0.18
1-phenyl-2-butanone	13.16	5.23	2.67	1.41	0.74	--
ethylbenzene	65.28	27.38	13.73	5.02	2.35	1.00
methyl phenylethyl ether	18.39	7.63	3.88	2.00	1.07	--
methyl 3-phenylpropionate	24.04	9.58	4.64	1.99	1.06	0.52
4-phenyl-2-butanone	12.60	5.24	2.42	1.35	0.75	--
2-phenylethanol	4.11	2.42	1.28	0.74	0.50	0.27
2-phenylethylamine	2.42	1.69	1.62	1.66	1.63	0.84
2-phenylethyl bromide	62.88	25.08	11.31	4.30	1.73	0.82
2-phenylethyl chloride	45.99	18.81	8.30	3.49	1.66	0.71
3-phenylpropionamide	2.65	1.38	0.83	0.51	0.36	0.23
3-phenyl-1-propionitrile	6.38	2.99	1.58	0.83	0.52	0.28
n-propylbenzene	165.50	59.57	24.60	8.19	3.38	1.29
methyl 4-phenylbutyrate	51.00	17.98	7.78	2.88	1.40	0.63
4-phenyl-1-butyronitrile	15.22	6.13	2.94	1.32	0.72	0.37
3-phenyl-1-propanol	8.55	3.84	2.06	1.03	0.62	0.31
3-phenyl-1-propylamine	4.80	3.15	2.73	2.43	2.35	--
3-phenyl-1-propyl bromide	163.37	56.07	22.12	7.21	2.92	1.10
3-phenyl-1-propyl chloride	121.95	42.36	15.69	5.86	2.45	0.95
n-butylbenzene	420.37	130.97	46.25	13.55	4.94	1.69

Table 4.5: Capacity factors of compounds with aliphatic substituents on an alkyl side chain in eluents containing acetonitrile

Compound	Capacity factor (*k*') Acetonitrile proportion (%)						
	30	40	50	60	70	80	90
toluene	28.86	12.74	6.08	3.24	1.97	1.23	0.75
benzyl alcohol	1.75	1.14	0.80	0.57	0.46	0.39	0.32
benzylamine	4.95	2.86	1.76	1.42	5.94	4.63	0.44
benzyl bromide	34.10	13.07	5.69	2.84	1.66	1.10	0.62
benzyl chloride	27.63	11.04	4.97	2.53	1.50	0.93	0.42
benzyl cyanide	6.87	3.50	1.92	1.13	0.75	0.52	0.35
methyl phenylacetate	10.63	4.99	2.59	1.47	0.95	0.65	0.43
phenylacetaldehyde	5.30	2.56	1.81	1.17	0.77	0.47	--
phenylacetamide	0.93	0.61	0.45	0.34	0.31	0.28	0.26
1-phenyl-2-butanone	15.38	5.89	3.24	1.85	1.10	0.62	--
ethylbenzene	59.53	22.38	9.46	4.66	2.67	1.58	0.91
methyl phenylethyl ether	16.39	6.38	3.76	2.19	1.37	0.79	--
methyl 3-phenylpropionate	20.39	8.26	3.85	2.04	1.26	0.81	0.53
4-phenyl-2-butanone	13.05	5.05	3.16	1.66	1.01	0.58	--
2-phenylethanol	2.71	1.56	1.00	0.68	0.53	0.43	0.35
2-phenylethylamine	1.45	1.36	1.41	1.47	7.54	7.81	0.44
2-phenylethyl bromide	61.35	20.70	8.18	3.83	2.13	1.23	0.72
2-phenylethyl chloride	47.24	16.74	6.85	3.28	1.86	1.09	0.61
3-phenylpropionamide	1.61	0.90	0.60	0.43	0.36	0.32	0.29
3-phenyl-1-propionitrile	9.47	4.40	2.27	1.27	0.82	0.55	0.38
n-propylbenzene	137.23	42.52	15.83	7.17	3.88	2.15	1.17
methyl 4-phenylbutyrate	38.09	13.35	5.63	2.79	1.64	1.01	0.62
4-phenyl-1-butyronitrile	18.40	7.38	3.41	1.78	1.08	0.69	0.46
3-phenyl-1-propanol	4.90	2.39	1.38	0.87	0.64	0.50	0.39
3-phenyl-1-propylamine	3.54	2.06	1.40	1.88	10.11	11.61	0.60
3-phenyl-1-propyl bromide	--	39.98	14.04	6.10	3.23	1.76	1.01
3-phenyl-1-propyl chloride	106.66	31.94	11.60	5.14	2.76	1.53	0.85
n-butylbenzene	308.86	79.97	26.33	11.02	5.63	2.94	1.51

Figure 4.2: Change in capacity factors of a selection of compounds with methanol concentration

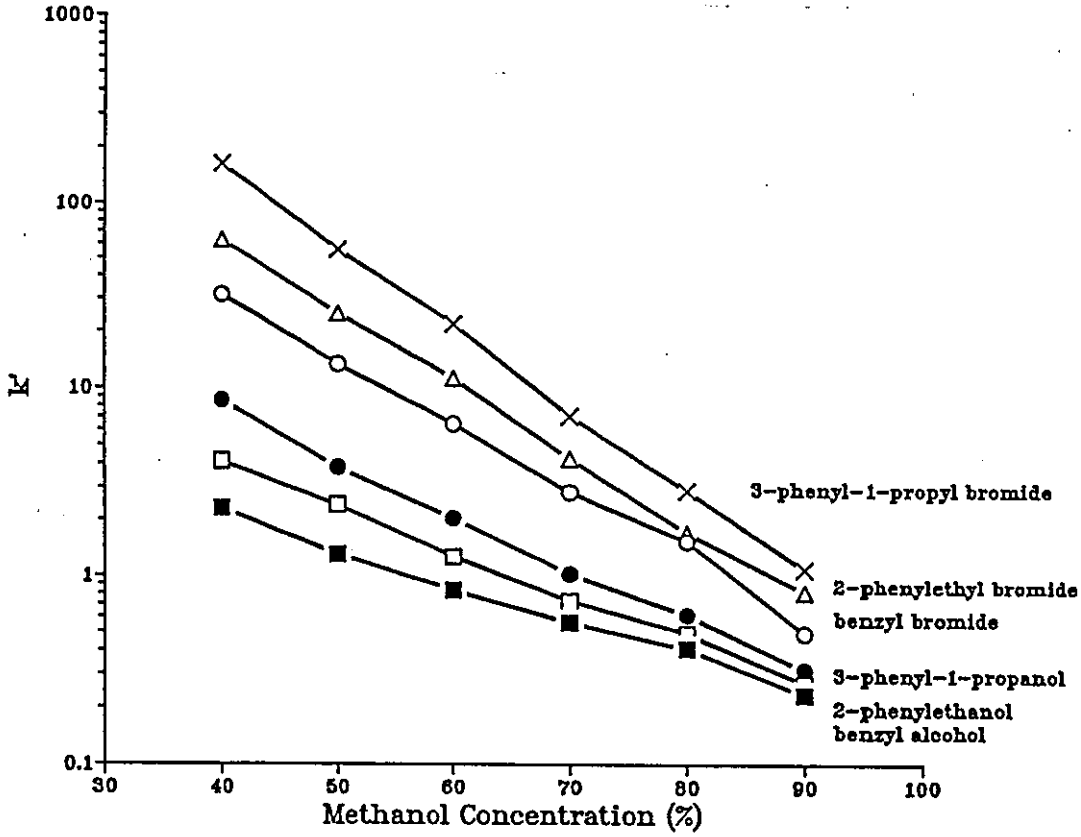
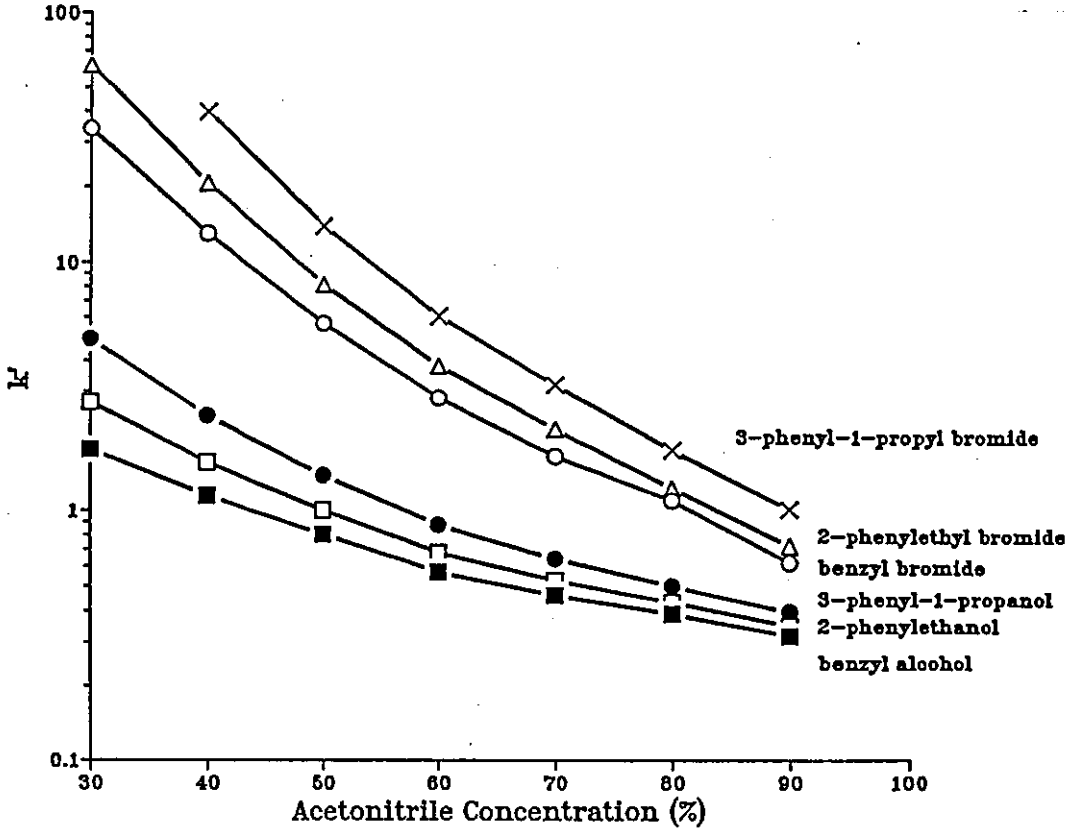


Figure 4.3: Change in capacity factors of a selection of compounds with acetonitrile concentration



Katz *et al.*[170] have suggested that at 80% methanol there is a change in selectivity due to changes in the nature of the mobile phase. A number of eluent modifier molecules would be aligned at the bonded phase surface and the change in the nature of this aligned "phase" would cause changes in the accessibility of the stationary phase for some compounds. It might be expected that this would be apparent in the log *k'* *vs.* eluent composition curves, however no such discontinuity was observed. The effects may be hidden as the capacity factors at 80% and 90% organic modifier were very small.

4.2.2 Retention Indices of Model Compounds

Retention indices of the model compounds (Tables 4.6 and 4.7) have been calculated from the linear relationship between log *k'* and $C_{n\infty}$ for the alkylarylketone standards. In each case the retention indices were determined using the calibration line derived from the set of alkylarylketones measured in the same run as the model compounds. As a consequence in these and equivalent tables in later chapters there may apparently be inconsistencies, in the relative order of elution, between the capacity factor and retention index tables. This comes about because capacity factors for different compounds, at particular eluent compositions may have been derived on different days and/or different columns.

For all the compounds the change in retention index was not as marked as the change in capacity factor. The retention indices of all the compounds showed some dependence on the eluent composition. Even the compounds

Table 4.6: Retention indices of aliphatic compounds in methanol eluents

Compound	Retention index Methanol proportion (%)					
	40	50	60	70	80	90
toluene	983	1010	1038	1063	1090	1120
benzyl alcohol	689	691	698	684	675	610
benzylamine	770	774	785	778	784	781
benzyl bromide	991	1004	1019	1030	1059	917
benzyl chloride	962	976	992	992	994	915
benzyl cyanide	773	766	763	738	722	658
methyl phenylacetate	863	862	853	853	846	837
phenylacetaldehyde	790	752	745	734	782	
phenylacetamide	627	623	591	598	582	519
1-phenyl-2-butanone	884	885	883	877	867	
ethylbenzene	1075	1100	1126	1151	1177	1204
methyl phenylethyl ether	923	934	945	955	968	
methyl 3-phenylpropionate	960	959	954	953	948	942
4-phenyl 2-butanone	888	883	879	871	863	
2-phenylethanol	756	774	759	741	729	667
2-phenethylamine	694	726	798	915	1073	1134
2-phenylethyl bromide	1072	1088	1107	1118	1090	1126
2-phenylethyl chloride	1033	1049	1066	1074	1077	1068
3-phenylpropionamide	705	698	667	663	640	602
3-phenyl-1-propionitrile	806	802	793	766	742	682
n-propylbenzene	1182	1204	1227	1256	1282	1308
methyl 4-phenylbutyrate	1046	1043	1040	1033	1029	1020
4-phenylbutyronitrile	907	899	892	866	842	799
3-phenyl-1-propanol	840	835	835	813	793	733
3-phenyl-1-propylamine	774	809	881	997	1178	
3-phenyl-1-propyl bromide	1181	1196	1214	1229	1240	1244
3-phenyl-1-propyl chloride	1147	1159	1175	1184	1190	1185
n-butylbenzene	1290	1310	1331	1364	1392	1419

Table 4.7: Retention indices of aliphatic compounds in acetonitrile eluents

Compound	Retention Index Acetonitrile proportion (%)						
	30	40	50	60	70	80	90
toluene	1005	1021	1033	1042	1050	1056	1067
benzyl alcohol	654	640	630	624	636	645	690
benzylamine	784	785	787	845	1366	1527	836
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
methyl phenylacetate	879	873	864	853	843	829	830
phenylacetaldehyde	781	784	781	790	790	782	-
phenylacetamide	575	540	516	504	522	535	599
1-phenyl-2-butanone	912	910	904	898	888	874	-
ethylbenzene	1095	1110	1121	1130	1137	1144	1153
methyl phenylethyl ether	919	925	932	937	945	957	-
methyl 3-phenylpropionate	961	952	942	932	922	907	915
4-phenyl-2-butanone	896	889	880	869	864	855	-
2-phenylethanol	708	689	675	667	676	682	730
2-phenethylamine	630	668	743	852	1434	1713	831
2-phenylethyl bromide	1099	1098	1092	1083	1073	1056	1053
2-phenylethyl chloride	1066	1064	1057	1045	1034	1013	975
3-phenylpropionamide	643	602	573	557	568	574	648
3-phenyl-1-propionitrile	865	853	837	818	799	771	767
n-propylbenzene	1200	1212	1223	1233	1244	1253	1264
methyl 4-phenylbutyrate	1039	1029	1018	1007	997	984	988
4-phenylbutyronitrile	948	935	918	898	879	851	851
3-phenyl-1-propanol	782	756	739	728	731	735	787
3-phenyl-1-propylamine	700	733	806	912	1518	1855	969
3-phenyl-1-propyl bromide		1202	1200	1194	1191	1182	1202
3-phenyl-1-propyl chloride	1168	1167	1162	1153	1147	1132	1126
n-butylbenzene	1302	1313	1325	1337	1350	1365	1377

which would be expected to belong to the same Snyder selectivity group[171] as the alkylarylketones, e.g. methyl 4-phenylbutyrate, show some changes in retention index across the eluent ranges. Smith and co-workers[134,144,145] have found that the largest changes in retention index were found with compounds which are considerably more polar than the alkylarylketone standards. The same effect was also found in this work with the polar compounds showing a large decrease in retention index with increasing modifier proportion and the non-polar compounds showing a gradual increase in retention index.

The changes in the retention indices in the eluents containing up to 80% organic were linear for only a few of the compounds (Figures 4.4 and 4.5). However, with the exception of the amines, the changes were systematic and could probably be described using quadratic regression equations (although this was not done).

For several compounds, benzyl alcohol, benzyl bromide, benzyl chloride, benzyl cyanide, phenylacetamide, and 3-phenyl-1-propanol, the changes between 80 and 90% organic are considerably larger than those observed over the rest of the eluent ranges. These observations suggest that the selectivity change which was described by Katz *et al.*[170], in eluents containing 90% methanol may be having an influence in the present study. Katz *et al.* reported that the change in selectivity was only significant in methanol eluents, however, relatively large changes in *RI* have been observed in both methanol and acetonitrile. At this modifier concentration the capacity factors were often small ($k' < 0.5$, Tables 4.4 and 4.5) and therefore very susceptible to uncertainties in the column void volume. The retention

Figure 4.4: Change in retention index of a selection of compounds with methanol concentration

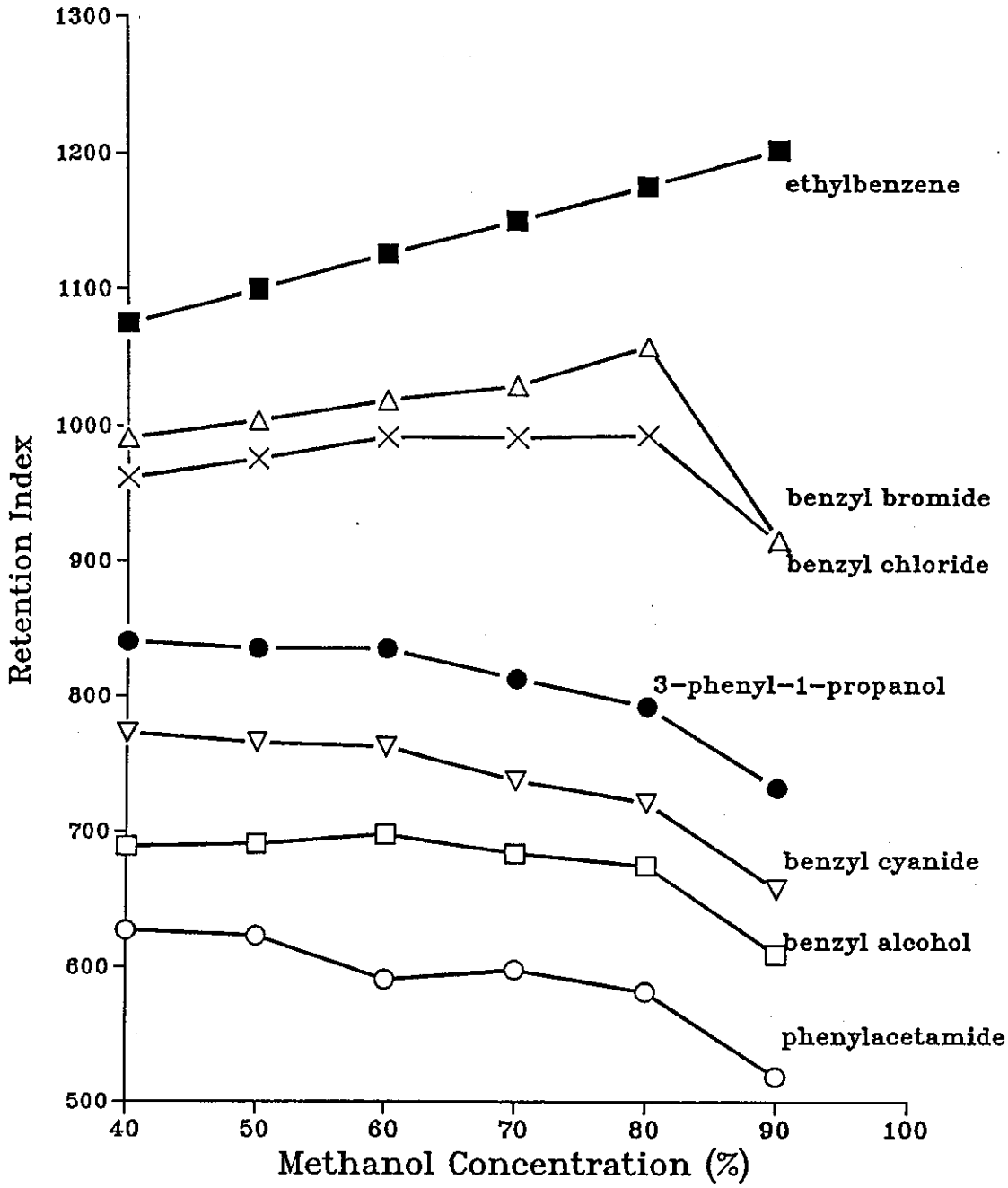
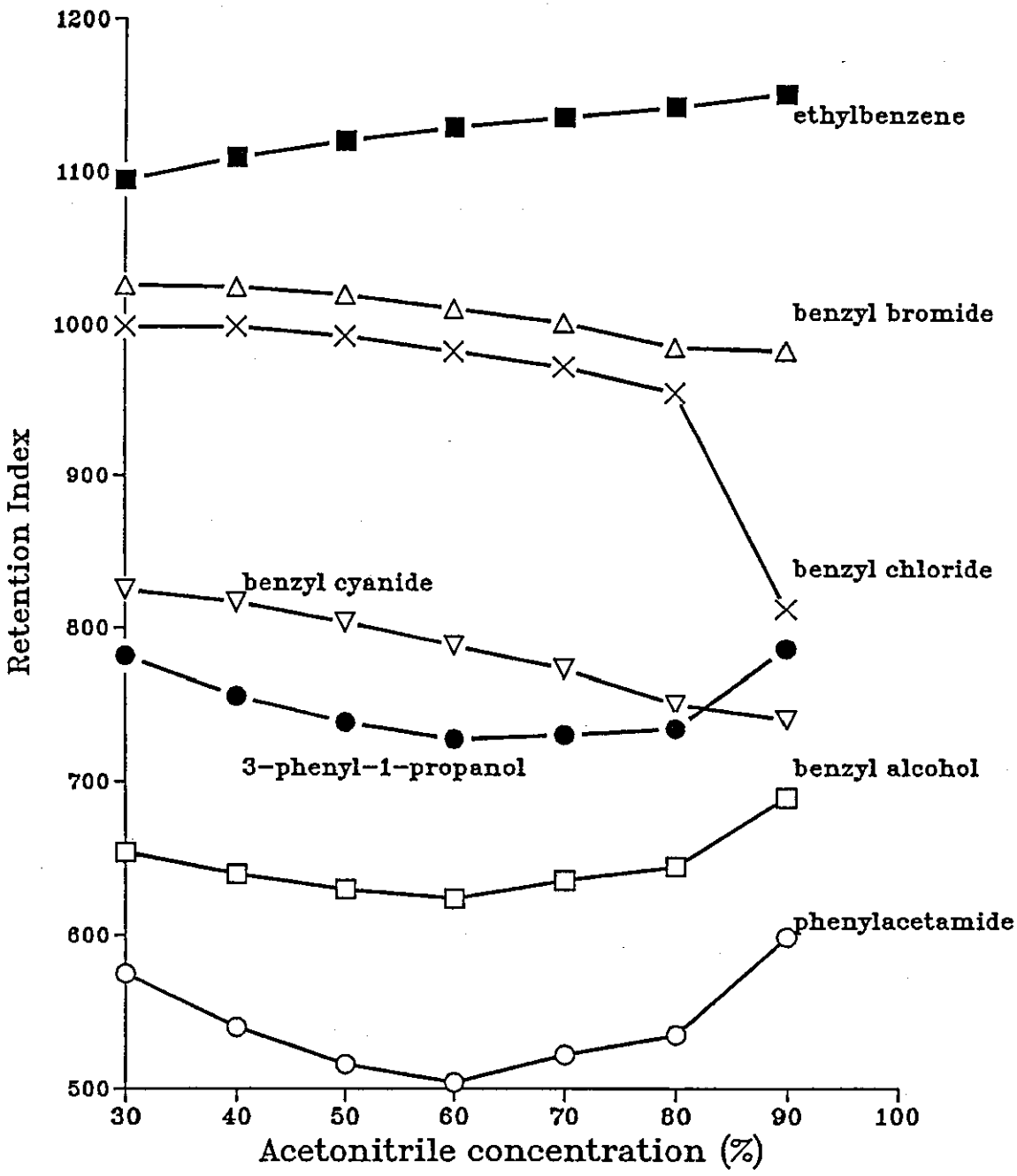


Figure 4.5: Change in retention index of a selection of compounds with acetonitrile concentration



indices have been shown to be more sensitive to the column void volume where the capacity factors were small (Chapter 3, Section 3.2.5). The relatively large changes in *RI* could therefore also be a reflection of the errors associated with the small *k'* values. Because of these marked changes in the retention indices and their uncertainty and corresponding changes seen later for some aromatic compounds (Chapter 6), it was decided to limit the the study eluent ranges up to 80% organic modifier. The validity of the prediction system was therefore also limited to the same ranges and as seen earlier the regression equation for benzene was only determined for this range.

4.2.3 Retention Behaviour of benzylamine, 2-phenylethylamine and 3-phenyl-1-propylamine

Benzylamine, 2-phenylethylamine and 3-phenyl-1-propylamine were not discussed in the previous sections as the changes in their capacity factors and retention indices across the eluent range were very different to the those of the other compounds. The capacity factors of these compounds (Table 4.4 and 4.5) were either virtually unchanged or increased with increasing organic modifier concentration (Figures 4.6 and 4.7). In methanol eluents the capacity factors of 2-phenylethylamine and 3-phenyl-1-propylamine were virtually constant between 50 and 80% modifier while in acetonitrile an increase was observed between 60 and 80% modifier. This suggests that the retention mechanism was probably more complicated than the usual reversed phase partition mechanism. The retention indices (Tables 4.6 and 4.7, Figures 4.8 and 4.9) also showed different patterns to

Figure 4.6: Change in capacity factors of benzylamine, 2-phenylethylamine and 3-phenyl-1-propylamine with methanol concentration

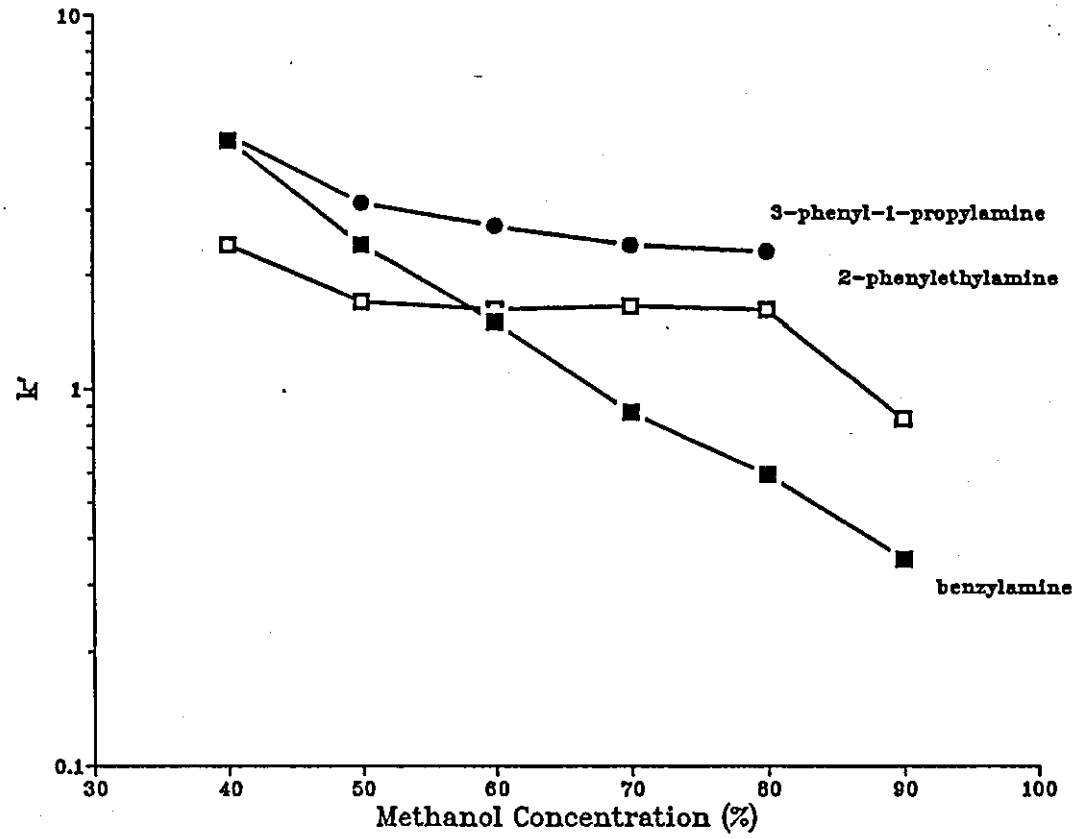
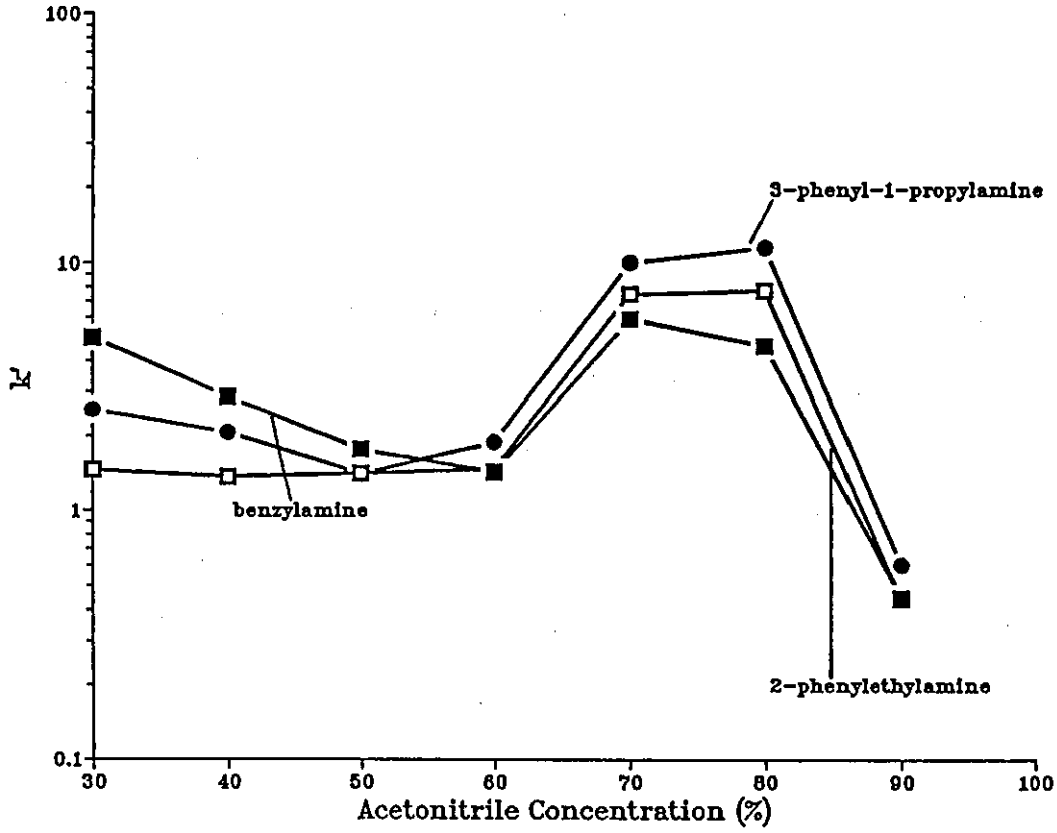


Figure 4.7: Change in capacity factors of benzylamine, 2-phenylethylamine and 3-phenyl-1-propylamine with acetonitrile concentration



those observed with other compounds. The retention index of 3-phenyl-1-propylamine increased from 700 in 30% acetonitrile to 1855 in 80% acetonitrile, in methanol the change was from 774 in 40% to 1178 in 80% and it was not eluted in 90% methanol.

Anomalous behaviour of basic compounds in RP-HPLC has been discussed by several workers and various patterns have been observed including changes which gave minima in the log k' against eluent composition relationship[164,172]. The observed behaviour was attributed to a mixed retention mechanism, of a combination of ion exchange chromatography and partition chromatography. The basic amines can interact strongly with the highly acidic silanol groups and are retained by an ion-exchange mechanism. The extent of this interaction depends on the nature of the stationary phase, in particular the amount of coverage i.e. the number of reacted and unreacted silanol groups[164]. Various workers have suggested the addition of competing amines to the mobile phase to preferentially bind to the silanols, this has succeeded in improving the peak shape and reproducibility of retention of basic compounds[173]. It should be noted that these compounds were not determined in a single injection so there will be no competition between the compounds for access to the silanol groups. In previous studies Smith *et al.*[149] have studied N-alkylanilines under conditions in which the eluent pH was controlled. In eluents of pH 8.2 the compounds would not be expected to be ionised and no abnormal behaviour was observed. The retention index was also not affected by the ionic strength of the mobile phase. However, in a study using conditions where the amines would be expected to be ionised the retention indices were

Figure 4.8: Change in retention indices of benzylamine, 2-phenylethylamine and 3-phenyl-1-propylamine with methanol concentration

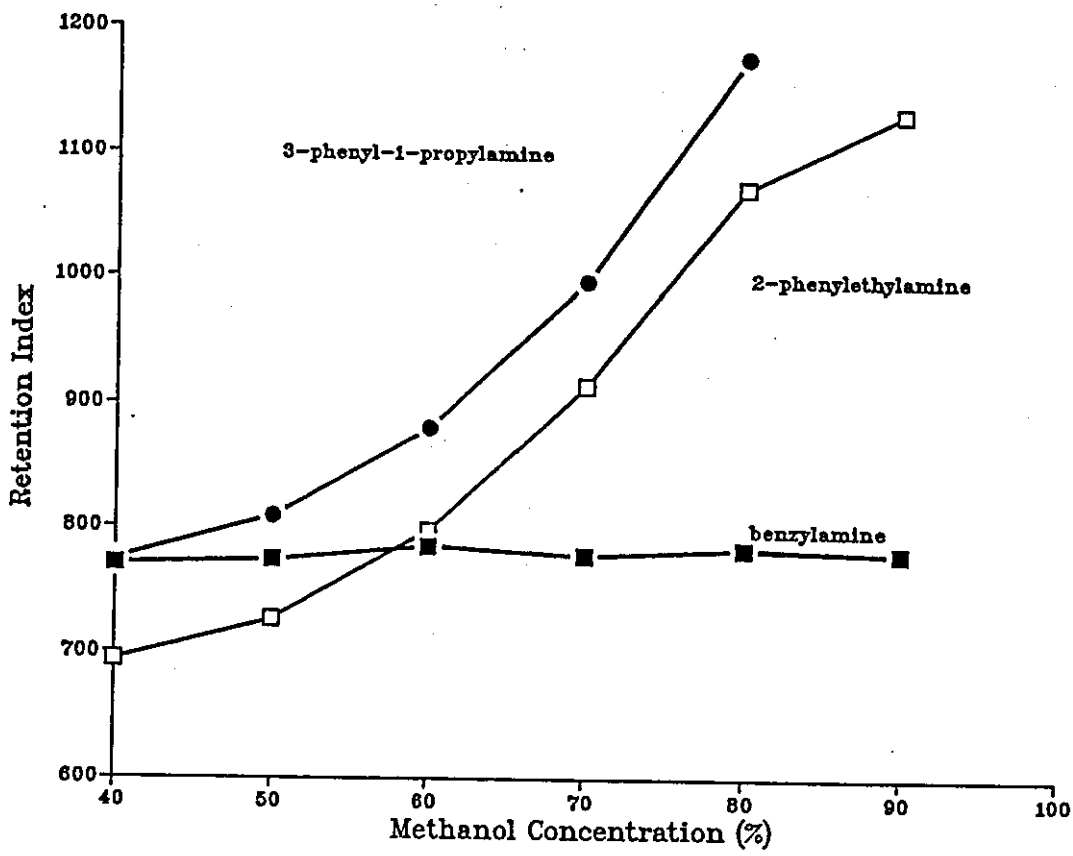
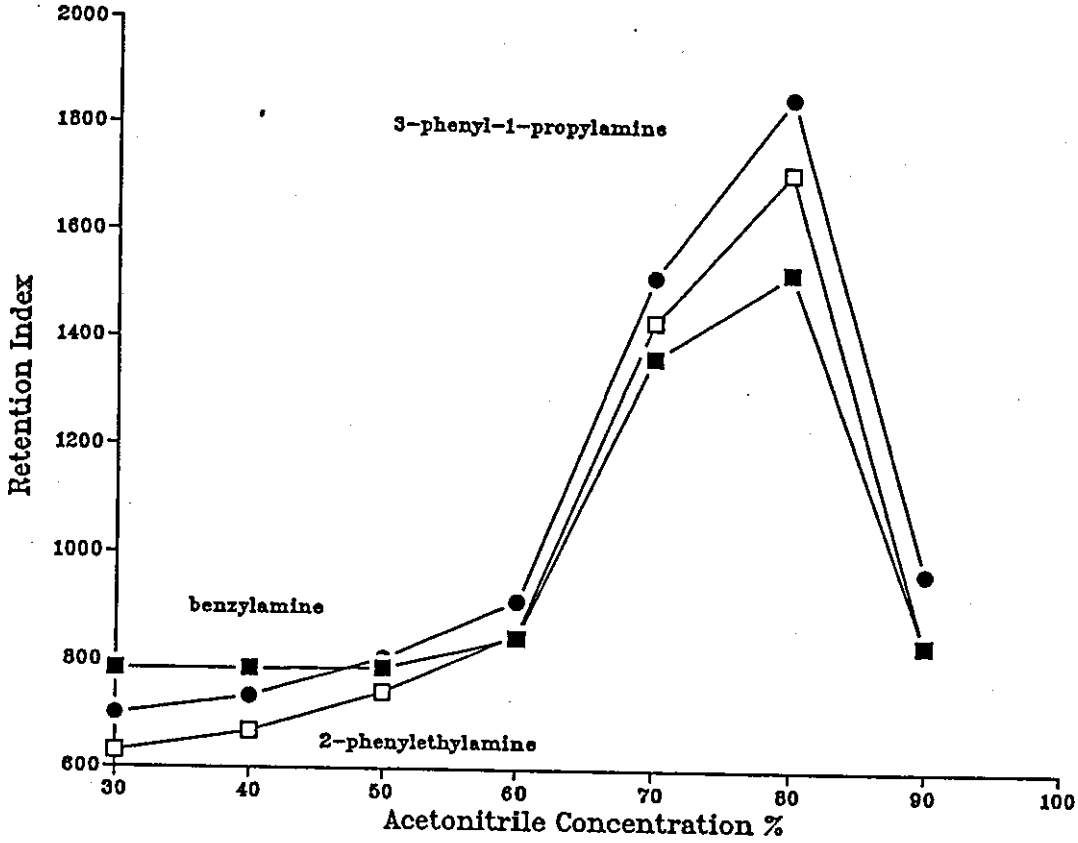


Figure 4.9: Change in retention indices of benzylamine, 2-phenylethylamine and 3-phenyl-1-propylamine with acetonitrile concentration



considerably more dependent on the eluent conditions even though an aliphatic amine was added to the eluent[149].

In the present case a form of ion-exchange retention mechanism was also supported by the study on the effect of ionic strength on the retention of 3-phenyl-1-propylamine (Chapter 3, Section 3.3). The retention indices and k' showed characteristic increases with decreasing the ionic strength of the eluent. The pKa values of the benzylamine, 2-phenylethylamine and 3-phenyl-1-propylamine are 9.35, 9.83 and 10.39[163] respectively, therefore under the conditions of the study (pH 7) they would be expected to be protonated to a large extent. The difference in pKa does not explain why benzyl amine behaves differently to the other amines across the methanol range and in low acetonitrile, eluents although the observations were probably related to the extent of ionisation.

An unusual effect was observed on going from 80 to 90% acetonitrile when the reduction in the ionic strength of the aqueous phase would be expected to increase the retention index, however, there was actually an observed reduction in the retention index.

As the retention indices of the amines would be very dependent on the eluent composition and the retention was not a simple partition process any substituent indices derived from this data would probably be highly irreproducible. The substituent indices would also probably not be applicable to other compounds where the degree of ionisation would differ. The retention indices of these compounds were not used to calculate substituent increments and substituent indices, it was not therefore possible to include a term for an aliphatic amino group into the

database used for retention prediction.

4.3 SUBSTITUENT INDICES FOR ALIPHATIC SUBSTITUENTS

The parent compound throughout the study was benzene (Section 4.1) and the retention index increments for the presence of substituents were calculated with reference to the smoothed parent index values rather than the experimental retention indices. As the substituents in this section were terminal substituents on an alkyl chain the retention index can be described using the relationship

$$RI_{Ph\langle CH_2\rangle n-x} = PI + nSI_{\langle CH_2\rangle} + SI_x \qquad 4.2$$

Before calculating the substituent indices the methylene group increment ($SI_{\langle CH_2\rangle}$)must therefore be known.

In the alkylarylketone homologous series this was defined as 100. The use of this value for the homologous alkylbenzene series has therefore been investigated.

4.3.1 Methylene Group Increments for Alkyl Side Chains

It has been suggested that the methylene group increments for different homologous series[4,174] may differ, however, this could lead to complications in the prediction system and the aim was therefore to establish whether the defined value of 100 could be used for other homologous series.

Using the parent index of benzene the retention increment for the alkyl side chain can be calculated using

the equation:-

$$\delta RI = RI_{PhC_nH_{2n+1}} - PI \qquad 4.3$$

The increment for a single methylene group can therefore be calculated as the difference between the retention indices of consecutive members of the homologous series (Table 4.8). Although the results all deviated from the defined value of 100, (the values ranged from 87 to 112), they were usually within the expected experimental uncertainty of +/- 10 (see Section 3.6). The exception was the increment obtained for increasing the chain length from toluene to that of ethylbenzene which had mean value in each modifier system of 88. Although there does not appear to be any reason for this methylene to be different from the subsequent carbons, a similar anomaly has also been observed in other physico-chemical properties. The octanol-water partition coefficient substituent constants (π)[49] of the alkyl chain fragments (Table 4.9) also showed an abnormal pattern. The increment between benzene and toluene was 0.56, the increment between ethylbenzene and n-propylbenzene was 0.53, and between n-propylbenzene and n-butylbenzene was 0.58. However the increment between toluene and ethylbenzene was 0.46.

In calculating the retention index increment for an alkyl side chain (SI_R) a correction of -12 (II_{PhCH-R}) for this abnormal behaviour of the second carbon has therefore been incorporated into the prediction system. For all other primary saturated carbon atoms the increment was taken as the defined value of 100.

Table 4.8: Retention index increments calculated for an alkyl chain

Members of series	Retention index increment Organic modifier proportion (%)					
	30	40	50	60	70	80
methanol-buffer						
toluene - benzene	--	98	97	100	102	108
ethylbenzene - toluene	--	92	90	88	88	87
propylbenzene- ethylbenzene	--	107	104	101	105	105
butylbenzene - propylbenzene	--	108	106	104	105	105
acetonitrile-buffer						
toluene - benzene	95	94	93	91	92	95
ethylbenzene - toluene	90	89	88	88	87	88
propylbenzene- ethylbenzene	105	102	102	103	107	109
butylbenzene - propylbenzene	102	101	103	107	106	112

Table 4.9: Octanol-water partition coefficient substituent constants (π) for alkyl side chains

From reference 48

Side chain	π.
CH_3	0.56
CH_2CH_3	1.02
$CH_2CH_2CH_3$	1.55
$CH_2CH_2CH_2CH_3$	2.13

4.3.2 Substituent Indices of Substituents on the Aliphatic Side Chain

The retention index increments (Tables 4.10 and 4.11) for the aliphatic substituents were calculated as the difference between the retention index of the model compound and the sum of the parent index and alkyl chain contribution at each eluent composition.

$$\delta RI = RI_{PhCnH2nX} - (PI + nSI_{CH2} + II_{PhCH2-R}) \quad 4.4$$

For all the substituents there were differences in the retention index increments calculated from the different parents, suggesting that there was some interaction with the aromatic ring in all cases. Any interactions, such as hyperconjugation or resonance effects, with the aromatic ring would be expected to be very short range and not significant after the alkyl chain was greater than about two carbon atoms. The substituent index equations (Table 4.12) have therefore been calculated using the retention index increments derived from the substituent on the longest available alkyl chain, in most cases this was for substitution on n-propylbenzene. The results suggest that there may be a problem with the substituent index equations for the aldehyde and ether groups which were based on a single example substituted on the benzylic carbon atom. The amide substituent index equation was also calculated from a substituent on a shorter alkyl chain. The substituent index equations for the aldehyde, amide and ether have been included in the database as the best information available but may need to be updated when examples in which the

Table 4.10: Retention index increments calculated in methanol eluents

Substituent	Retention index increment Methanol proportion (%) 40	50	60	70	80
Substituted toluene					
Parent *RI* ($PI + SI_{CH2}$)	985	1013	1039	1061	1082
$CONH_2$	-358	-390	-447	-463	-500
OH	-296	-322	-340	-377	-407
CN	-212	-247	-275	-323	-360
*CHO	-195	-261	-293	-327	-300
CO_2CH_3	-122	-151	-185	-208	-236
(CO_2R[a]	-222	-251	-285	-308	-336)
COC_2H_5	-101	-128	-155	-184	-215
(COR[b]	-301	-328	-355	-384	-415)
Cl	-23	-37	-46	-69	-88
Br	6	-9	-19	-31	-23
Substituted ethylbenzene					
Parent *RI* ($PI + 2SI_{CH2} + II_{PhCH2-R}$)	1073	1101	1126	1149	1170
*$CONH_2$	-368	-403	-459	-486	-530
OH	-317	-327	-367	-408	-441
CN	-267	-299	-333	-383	-428
$COCH_3$	-189	-216	-243	-272	-303
*(COR[a]	-289	-316	-343	-372	-403)
OCH_3	-150	-167	-181	-194	-202
*(OR[a]	-250	-267	-281	-294	-302)
CO_2CH_3	-113	-142	-172	-196	-222
(CO_2R[a]	-213	-242	-272	-296	-322)
Cl	-40	-52	-60	-75	-93
Br	-1	-13	-19	-31	-80
Substituted propylbenzene					
Parent *RI* ($PI + 3SI_{CH2} + II_{PhCH2-R}$)	1173	1201	1226	1249	1270
*OH	-333	-366	-391	-436	-477
*CN	-266	-302	-334	-383	-428
CO_2CH_3	-127	-158	-186	-216	-241
*(CO_2R[a]	-227	-258	-286	-316	-341)
*Cl	-26	-42	-51	-65	-80
*Br	8	-5	-12	-20	-70

* values used for the calculation of substituent index equations
[a] 100 subtracted for the methyl group contribution
[b] 200 subtracted for the ethyl group contribution

Table 4.11: Retention index increments calculated for aliphatic substituents in eluents containing acetonitrile

Substituent	Retention index increment Acetonitrile proportion (%) 30	40	50	60	70	80
Substituted toluene						
Parent *RI* ($PI + SI_{CH2}$)	1001	1027	1040	1051	1058	1063
$CONH_2$	-435	-487	-524	-547	-536	-528
OH	-356	-387	-410	-427	-422	-418
CN	-185	-210	-236	-262	-284	-312
*CHO	-229	-243	-259	-261	-268	-281
CO_2CH_3	-131	-154	-176	-198	-215	-234
(CO_2R[a]	-231	-254	-276	-298	-315	-334)
COC_2H_5	-98	-117	-136	-153	-170	-189
(COR[b]	-298	-317	-336	-353	-370	-389)
Cl	-11	-28	-47	-68	-85	-107
Br	16	-2	-20	-40	-56	-77
Substituted ethylbenzene						
Parent *RI* ($PI + 2SI_{CH2} + II_{PhCH2-R}$)	1098	1115	1128	1139	1146	1151
*$CONH_2$	-455	-513	-555	-582	-578	-577
OH	-390	-426	-453	-472	-470	-469
CN	-233	-262	-291	-321	-347	-380
$COCH_3$	-202	-226	-248	-270	-282	-296
*(COR[a]	-302	-326	-348	-370	-382	-396)
OCH_3	-179	-190	-196	-202	-201	-194
*(OR[a]	-279	-290	-296	-302	-301	-294)
CO_2CH_3	-137	-163	-186	-207	-224	-244
(CO_2R[a]	-237	-263	-286	-307	-324	-344)
Cl	-32	-51	-71	-94	-112	-138
Br	1	-17	-36	-56	-73	-95
Substituted propylbenzene						
Parent *RI* ($PI + 3SI_{CH2} + II_{PhCH2-R}$)	1198	1215	1228	1239	1246	1251
*OH	-416	-459	-489	-511	-515	-516
*CN	-250	-280	-310	-341	-367	-400
CO_2CH_3	-159	-186	-210	-232	-249	-267
*(CO_2R[a]	-259	-286	-310	-332	-349	-367)
*Cl	-30	-48	-66	-86	-99	-119
*Br	-	-13	-28	-45	-55	-69

* values used for the calculation of substituent index equations
[a] 100 subtracted for the methyl group contribution
[b] 200 subtracted for the ethyl group contribution

substituent is on a longer alkyl chain can be examined.

A number of the functional groups, CO_2CH_3, OCH_3, and $COCH_3$, contain saturated alkyl groups not directly attached to the aromatic ring. In these compounds the retention index increments for the alkyl group have been calculated based on the definition of the methylene increment equal to 100 and can be subtracted from the value of the whole group. The values for the functional groups CO_2R, OR and COR have therefore been listed in the tables and these values were used to calculate the substituent index equations.

For all the substituents the fitted quadratic equations were a good description of the experimentally determined retention index increments (Figures 4.10 and 4.11, the points were the experimentally determined retention index increments and the lines the fitted regression equations).

The intercept (c) values in the two modifiers should be equal to the substituent index at 100% water and therefore the same in both methanol and acetonitrile. However, considerable differences were observed, for example in methanol the intercept for the bromide was -99 and in acetonitrile 63, for the aldehyde the intercept was +337 in methanol and -184 in acetonitrile. These differences emphasise the importance of not extrapolating the quadratic equations outside of the calibrated region. The coefficients of the fitted equations were also found to be very sensitive to the small changes in the values used in their calculation.

It was interesting to note that the coefficients for the aldehyde group (CHO), which should be the first member of the series COR, and the ketone substituent (COR) differed greatly in both eluents although the contribution of a

Table 4.12: Substituent index equations for aliphatic substituents on an alkyl side chain

$$SI = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Substituent	Coefficients of substituent index equations		
	a	b	c
methanol-buffer			
$CONH_2$	0.0079	-5.013	-178
OH	-0.0257	-0.494	-273
CN	-0.0250	-1.050	-185
CHO	0.1314	-18.531	337
CO_2R	0.0071	-3.717	-90
COR	-0.0086	-1.851	-201
OR	0.0136	-2.939	-154
Cl	-0.0021	-1.053	19
Br	-0.0536	4.719	-99
acetonitrile-buffer			
$CONH_2$	0.0855	-11.786	-179
OH	0.0561	-8.139	-223
CN	0.0002	-2.997	-160
CHO	0.0073	-1.768	-184
CO_2R	0.0130	-3.580	-164
COR	0.0161	-3.654	-206
OR	0.0211	-2.644	-218
Cl	0.0018	-1.962	27
Br	0.0064	-2.161	63

Figure 4.10: Relationship between substituent index calculated from regression equation and substituent increment for a selection of aliphatic substituents in methanol eluents

Points are experimental retention increments and the lines the calculated substituent indices

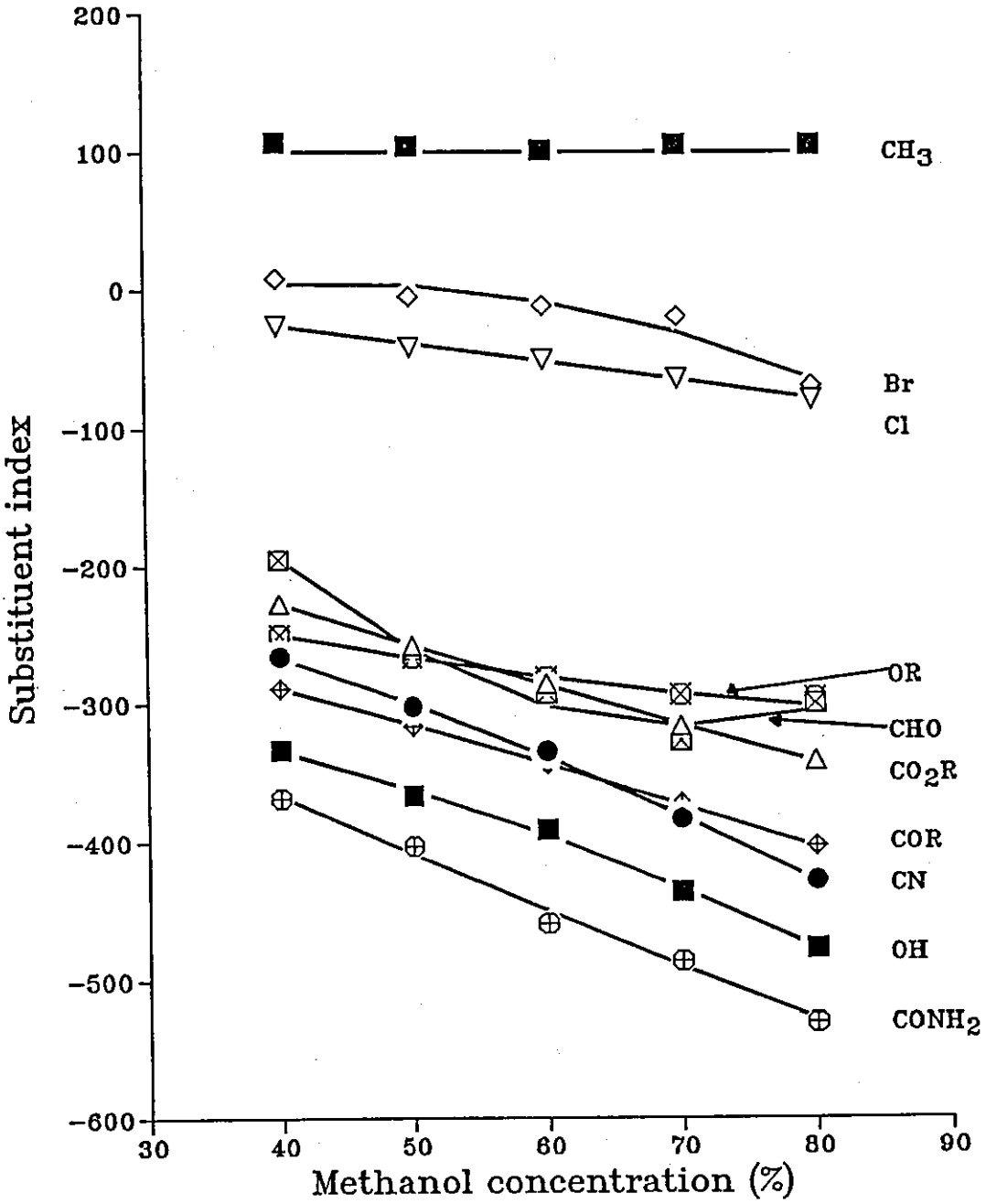
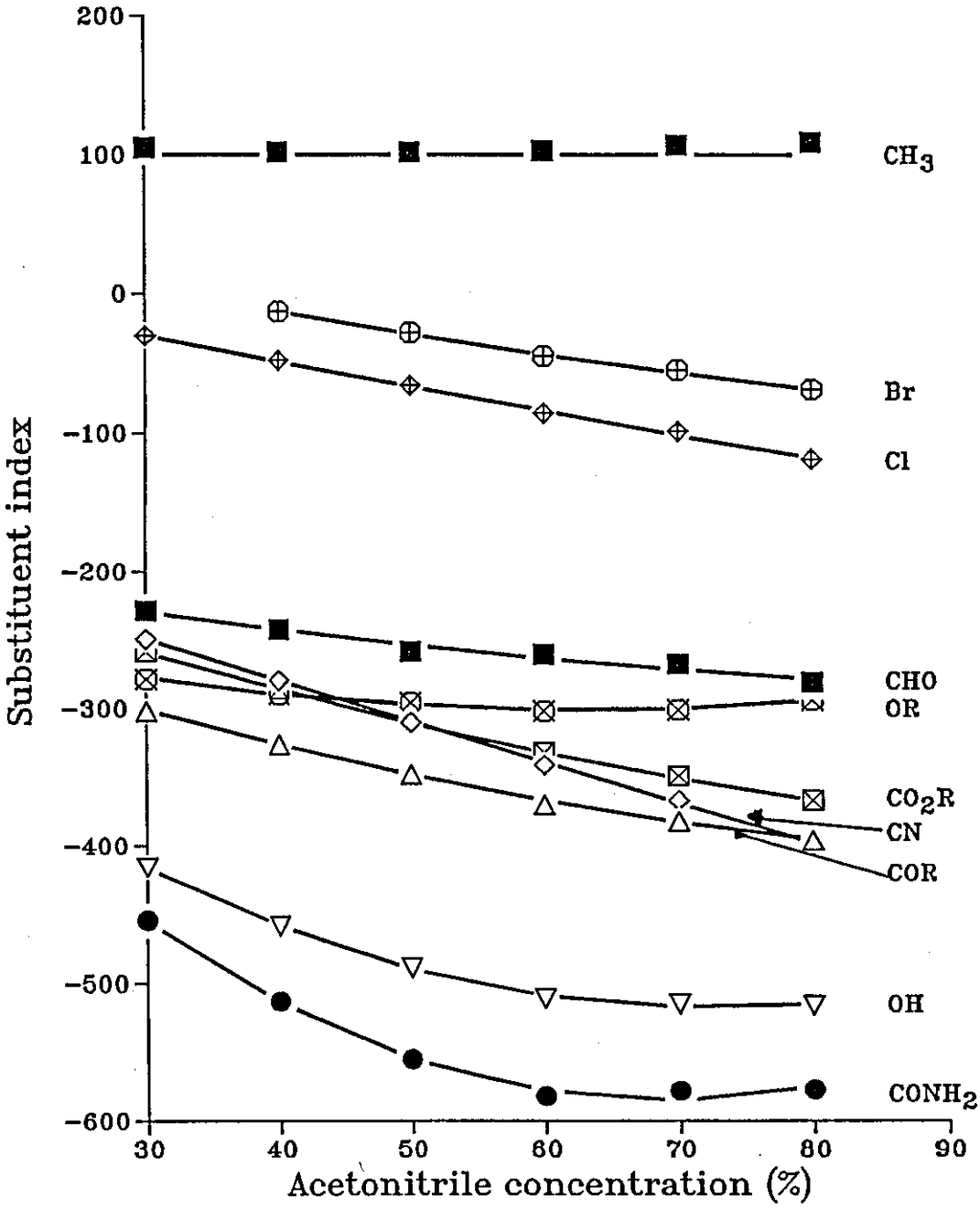


Figure 4.11: Relationship between substituent index calculated from regression equation and substituent increment for a selection of aliphatic substituents in acetonitrile eluents

Points are experimental retention increments and the lines the calculated substituent indices



hydrogen to the retention index was defined as 0. This may at least partially be due to the use of values derived from different lengths of homologous series, however, the retention indices of phenylacetaldehyde and 1-phenyl-2-butanone did not differ by the 200 expected but by only just over 100.

These substituent index equations have been included in the database and can be used for the prediction of retention indices of other compounds (Chapter 8).

4.4 EFFECT OF UNSATURATION IN THE ALKYL CHAIN ON RETENTION INDEX

The effect of introducing a double bond into an alkyl chain has been examined using two compounds 1-phenyl-1-propene ($PhCH{=}CHCH_3$) and 3-phenyl-1-propene ($PhCH_2CH{=}CH_2$). These compounds differed only in the position of the double bond. The capacity factors of these compounds (Table 4.13) were used to calculate retention indices (Table 4.14). If the double bond made no contribution to the retention of these compounds the side chain would be expected to contribute 288 to the retention index in both eluents.

$$SI_{chain} = 3 \times SI_{CH_2} + II_{PhCH_2-R} = 288 \qquad 4.5$$

In practice the retention index increment due to the addition of the alkyl chain differed considerably from this value (Table 4.15) with the double bond causing a significant reduction in the retention compared to the unsaturated compound. The unsaturated chain was apparently

Table 4.13: Capacity factors of 3-phenyl-1-propene and 1-phenyl-1-propene

Compound	Capacity factor Modifier proportion (%) 30	40	50	60	70	80
methanol-buffer						
3-phenyl-1-propene		91.64	31.68	12.03	5.41	2.39
1-phenyl-1-propene		112.43	38.39	14.85	6.17	2.80
acetonitrile-buffer						
3-phenyl-1-propene	83.52	29.08	10.38	5.06	2.63	1.28
1-phenyl-1-propene	96.47	32.77	13.13	5.56	2.90	1.43

Table 4.14: Retention indices of 3-phenyl-1-propene and 1-phenyl-1-propene

Compound	Retention index Modifier proportion (%) 30	40	50	60	70	80
methanol-buffer						
3-phenyl-1-propene		1105	1133	1151	1166	1186
1-phenyl-1-propene		1128	1156	1176	1194	1220
acetonitrile-buffer						
3-phenyl-1-propene	1119	1127	1132	1135	1135	1138
3-phenyl-2-propene	1136	1144	1149	1157	1163	1175

more polar than the saturated chain. The effect on retention was dependent on the position of the double bond relative to the ring (Table 4.15). Conjugation with the aromatic ring increasing the retention index increment significantly. The two functional groups were therefore treated separately as an aliphatic unsaturated functional group (PhRCH:CHR) and an aromatic unsaturated functional group (PhCH:CHR). In each case the change across the eluent range was described by a quadratic regression equation (Table 4.16).

4.5 RELATIONSHIP BETWEEN *SI* AND STRUCTURAL PARAMETERS

It was intended to examine whether the substituent index values calculated from the retention indices could be related to physico-chemical properties of the substituents.

4.5.1 Relationship between *SI* and π

The relationship between octanol-water partition coefficient (log *P*) and retention in RP-HPLC was discussed earlier. Briefly retention in RP-HPLC is governed by the solvophobic mechanism proposed by Horvath *et al.*[2,3]. In this retention mechanism the role of the stationary phase is seen as being that of a passive receptor with the mobile phase interactions playing a predominant role. The retention mechanism is therefore a liquid-liquid partition between a non-polar bonded phase and the polar eluent. The most common liquid-liquid partition values available are the octanol-water partition coefficient which are frequently used to model quantitative structure-activity relationships (QSAR)

Table 4.15: Retention index increments for alkene chains derived from 3-phenyl-1-propene and 1-phenyl-1-propene

Calculated from the equation $\delta RI = RI - PI$

Group	Retention index increment Modifier proportion (%) 30	40	50	60	70	80
methanol-buffer						
$CH_2CH:CH$		222	218	214	208	183
(RCH:CH*		122	118	114	108	83)
$CH:CHCH_2$		245	241	239	236	211
(CH:CHR*		145	141	139	136	111)
acetonitrile-buffer						
$CH_2CH:CH$	209	200	192	184	175	178
(RCH:CH*	109	100	92	84	75	78)
$CH:CHCH_2$	226	217	209	206	203	215
(CH:CHR*	126	117	109	106	103	115)

* calculated as the effect of CH:CH by subtraction of 100 for the saturated methyl/methylene group

Table 4.16: Substituent index equations for aromatic alkene group and aliphatic alkene group

$$SI = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Functional group	Coefficients of substituent index equations a	b	c
methanol - buffer			
aromatic CH:CH	-0.0307	2.956	74
aliphatic CH:CH	-0.0314	2.891	55
acetonitrile - buffer			
aromatic CH:CH	0.0223	-2.741	189
aliphatic CH:CH	0.0100	-1.780	154

studies in drug compounds. Many workers have found an approximately linear relationship between log *P* and log *k'* for related groups of compounds[1,55,175]. Hansch[48] and Rekker[53] have both proposed methods of calculating log *P* as the sum of the contributions of the substituents and the parent. Of the two methods proposed the Hansch approach has probably been more frequently used and a large number of substituent contributions (π) have been tabulated[48]. If a linear relationship between π and *SI* exists it could be useful for several reasons. Firstly, it would confirm that the predominant retention mechanism is a liquid-liquid partition and secondly, it would highlight any functional groups for which a more complex mechanism is present which might cause problems in prediction. In addition if a linear relationship could be found it would be provide a method to estimate *SI* for a substituent not included in the original data set, e.g. un-ionised aliphatic amines and carboxylic acids.

The group contributions to octanol-water partition coefficients (π) (Table 4.17) have been used to obtain linear regression equations between π and substituent indices. The correlations coefficients obtained (Table 4.18) show that there was an approximately linear relationship between *SI* and π with the correlations ranging from 0.925 to 0.956 in methanol and 0.918 to 0.954 in acetonitrile. In neither methanol (Figure 4.12, for 40% methanol) or acetonitrile (Figure 4.13, for 40% acetonitrile) can any definite outliers be identified. Although in both cases the OH has a substituent index considerably more negative than would be predicted from the Hansch substituent constant. Various workers have suggested that hydrogen bonding and

non-hydrogen bonding species should treated separately when correlations of log *P* against log *k'* were considered[56,57].

There did not appear to be any significant improvement in the correlation coefficient with decreasing organic modifier concentration. This contrasted with previous reports which have suggested that the HPLC system more closely resembles that of octanol-water when the organic content was lowest[54]. The correlation between *SI* and π was lower in acetonitrile than methanol suggesting that the octanol-water partition coefficient was a better descriptor of the processes occurring in methanol than acetonitrile, Braumann[54] also suggested that this could be the case. From this work it would appear that the use of π would at best provide only a very rough estimate of the *SI* for a given substituent.

4.5.2 Relationship between *SI* and aqueous solubility

An alternative structural property which has been used to describe retention in RP HPLC is the aqueous solubility. Hafkenscheid *et al.*[80-83] found a linear relationship between log *k'* and the Hildebrand solubility parameter. This parameter is a measure of the solubility of the whole molecule. There were difficulties in using the relationship because of problems in either obtaining reliable published data or calculating the values from thermodynamic properties. Wakita *et al.*[176] have recently proposed a method of calculating aqueous solubility as the sum of the contributions of the different substituents. The derived f_w. values (substituent contributions to solubility) have been calculated for a number of aliphatic and aromatic

Table 4.17: Hansch (π) values and Wakita solubility fragment (f_m) for aliphatic substituents

Substituent	π[48]	R group	f_m[176]
CO_2R	(-0.64)	(CH_3)	-1.90
OH	-1.12		-2.65
CHO	-0.91		-1.82
OR	(0.03)	(CH_2CH_3)	-2.51
Cl	0.39		-0.12
Br	0.60		0.05
CN	-0.84		-1.29
COR	(-0.62)	(CH_3)	-2.43
CH_3	0.5		0.73
$CONH_2$	-1.71		-

the values in brackets are the alkyl substituents (i.e. R) for which the π value was available in the reference 48

Table 4.18 Regression equations obtained for least squares fit of *SI* vs. π.

$$SI = a\pi + b$$

Eluent	Coefficients of regression equation a	b	Correlation coefficient
methanol-buffer			
40:60	-65.85	183	0.9470
50:50	-83.31	203	0.9563
60:40	-101.58	218	0.9537
70:30	-120.64	229	0.9447
80:20	-140.83	237	0.9250
acetonitrile-buffer			
30:70	-77.90	217	0.9540
40:60	-95.35	233	0.9466
50:50	-110.80	243	0.9397
60:40	-119.78	251	0.9383
70:30	-136.53	248	0.9270
80:20	-146.11	243	0.9183

Figure 4.12: Relationship between substituent index and substituent π values in methanol eluents

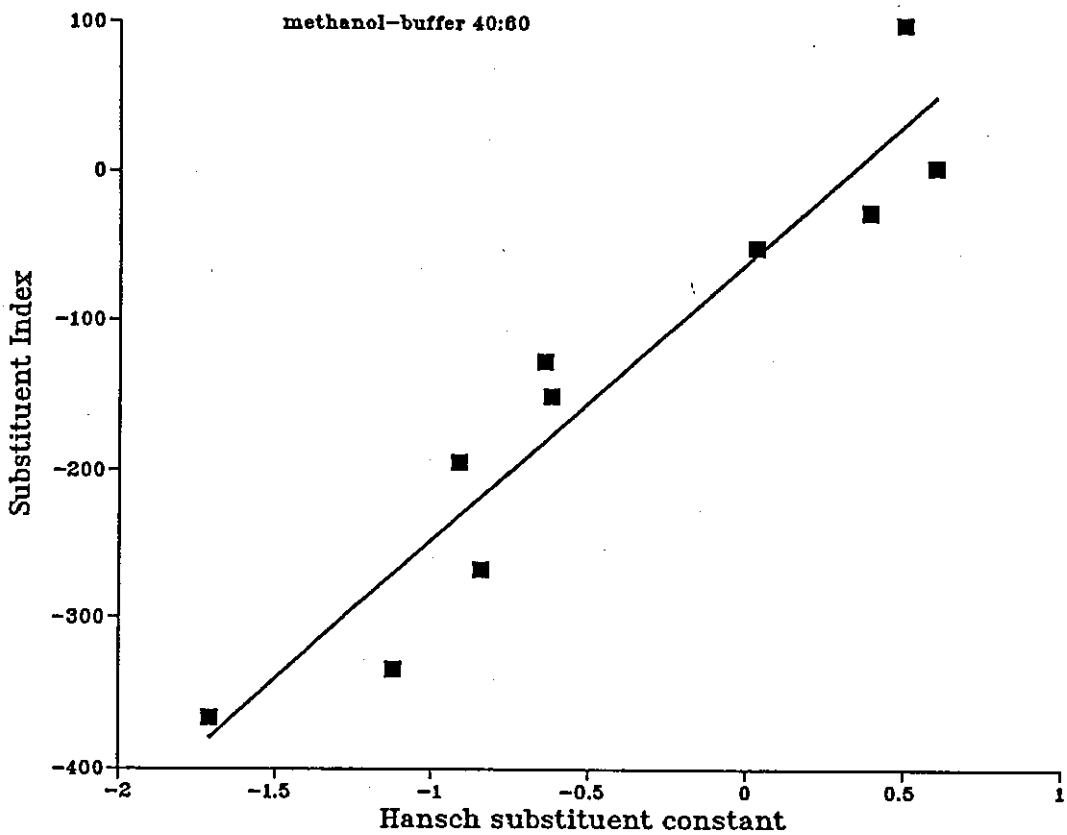
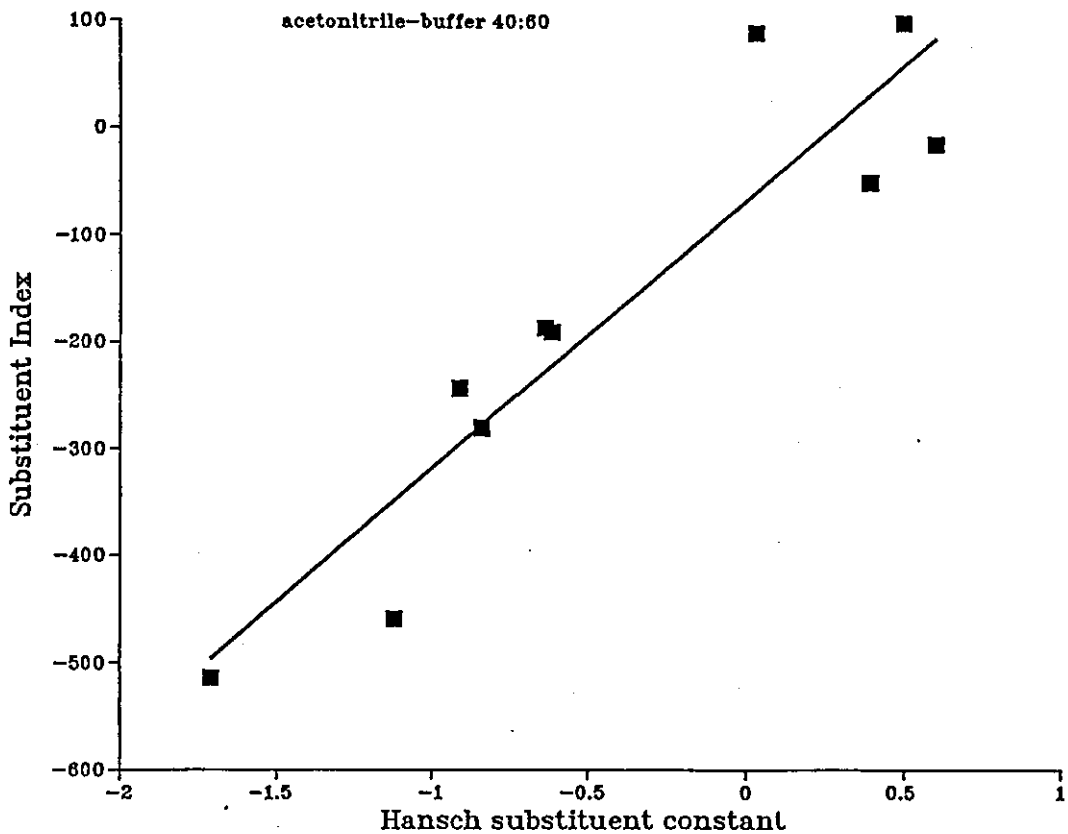


Figure 4.13: Relationship between substituent index and substituent π values in acetonitrile eluents



substituents (Table 4.17). In an manner analogous to that using π the correlation between *SI* and f_m has been calculated using a linear least squares regression equation (Table 4.19).

As with the π values the correlations showed a linear trend in the relationship but the correlations were again not high. As the f_m values were 1/solubility this suggests an increase in retention index with a decrease in water solubility. In the methanol eluents a single outlier was identifiable (Figure 4.14 in 40% methanol), this was the CN group. It appears that the behaviour of OH was well described by the water solubility probably because the method more accurately portrayed the interactions occurring than the octanol-water coefficients. In acetonitrile eluents there was an increase in linearity with decreasing acetonitrile content, the nitrile group also appears to be an outlier in this modifier system (Figure 4.15 at 40% MeCN).

Neither log *P* or aqueous solubility was a perfect descriptor of the processes occurring in RP-HPLC, however, they could prove useful for identifying substituents where the retention process was more complex, e.g. nitrile, and possibly where the assumption of a simple partition process was not valid. The correlations were not high but it might be possible to use the derived equations to estimate the *SI* for a substituent for which the *SI* was unknown but the confidence in the calculation would be low.

Table 4.19: Regression equations obtained for least squares fit of *SI* vs. f_m

$$SI = af_m + b$$

Eluent	Coefficients of regression equation a	b	Correlation coefficient
methanol-buffer			
40:60	-10.57	116	0.9580
50:50	-21.95	126	0.9552
60:40	-34.94	133	0.9456
70:30	-49.25	139	0.9341
80:20	-65.82	143	0.9148
acetonitrile-buffer			
30:70	-4.66	129	0.9700
40:60	-23.97	137	0.9597
50:50	-37.67	140	0.9493
60:40	-49.99	142	0.9371
70:30	-61.29	141	0.9257
80:20	-42.49	154	0.9095

Figure 4.14: Relationship between substituent index and substituent aqueous solubility constant in methanol

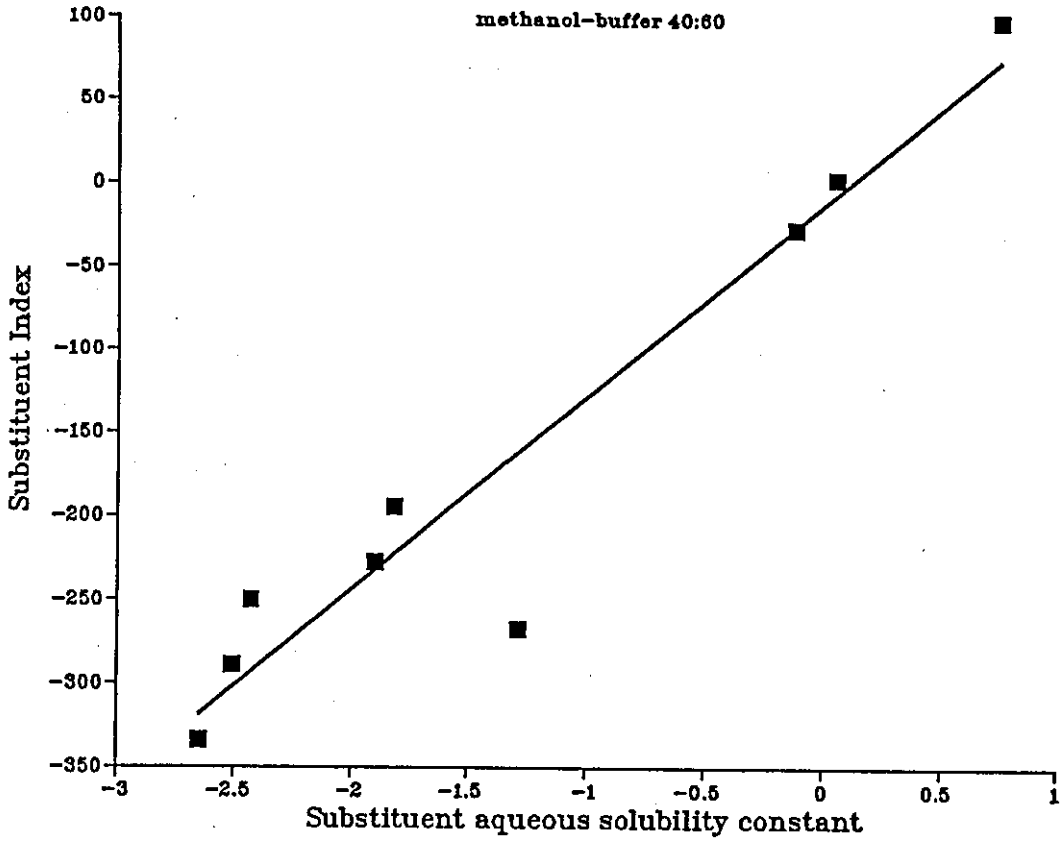
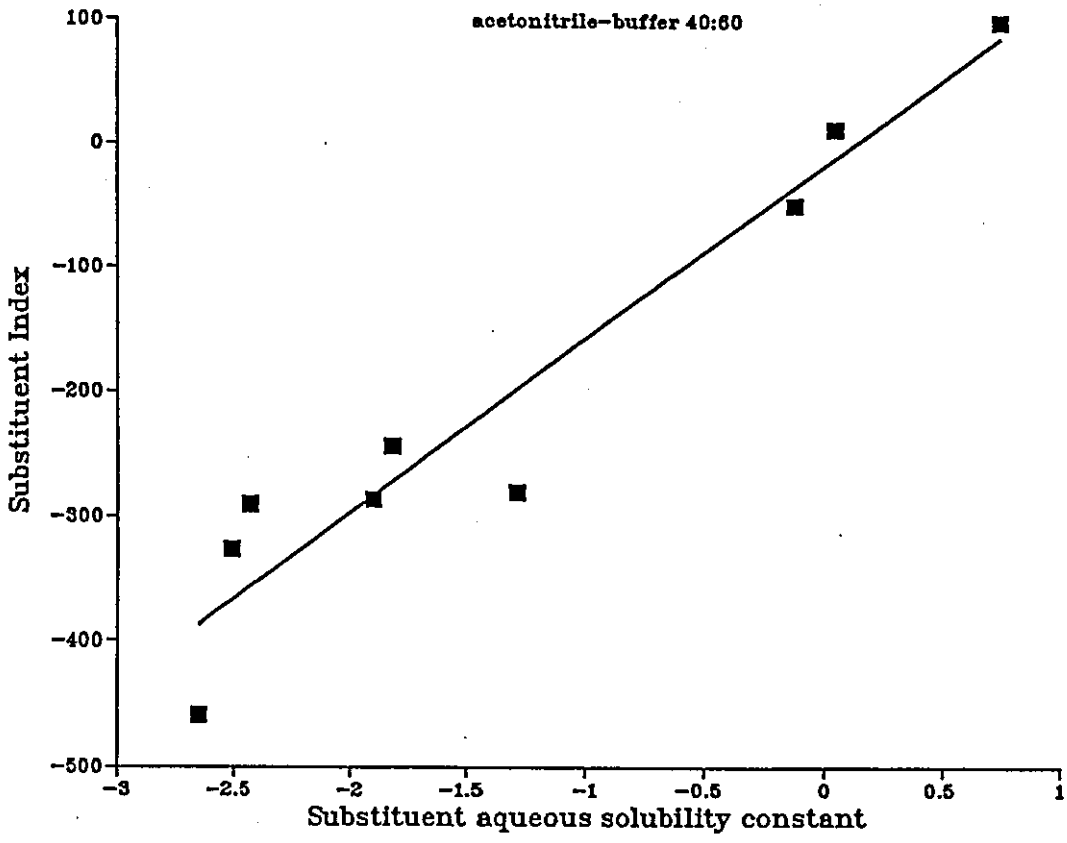


Figure 4.15: Relationship between substituent index and substituent aqueous solubility constant in acetonitrile



4.6 INTERACTIONS OF SUBSTITUENTS WITH THE AROMATIC RING

Where possible δRI values were calculated for functional groups on more than one length of alkyl chain enabling of any interactions between the substituent and ring to be identified. Any interactions between the substituents and the aromatic ring would be expected to depend on the type of substituent present, the distance from the ring and also the eluent composition. As the interactions would be expected to be electronic, for example electron withdrawing effects, they would be expected to be stronger with substituents positioned on the benzylic carbon. The size of the interactions were expressed as interaction indices (II). Using these values the retention index of an aliphatic substituted compound would be calculated using the equation

$$RI = PI + SI_x + nSI_{(CH_2)} + II_{PhCH_2x} \qquad 4.6$$

The effect of the aromatic ring would also be expected to be fairly short range and consequently not significant at C_{no} > about 2. The extent of the interaction (Table 4.20 and 4.21) has been examined by calculating the difference between the substituent index derived from the longest chain substituted compounds (from Table 4.12) and the retention index increment (from Tables 4.11 and 4.12) of the other substituents (Table 4.20 and 4.21) using equation 4.7.

$$\delta I = \delta RI_x - SI_x \qquad 4.7$$

As the aldehyde and ether groups were only examined on a

Table 4.20: Interaction increments for interaction of aliphatic substituents with the aromatic ring in methanol eluents

Substituent	Interaction increment Methanol proportion (%) 40	50	60	70	80
$\delta I = \delta RI_{PhCH2-x} - SI$					
$CONH_2$	11	19	3	27	28
OH	38	40	55	67	70
CN	55	53	53	58	69
CO_2R	5	7	2	7	6
COR	-12	-13	-12	-13	-12
Cl	3	2	5	-4	-9
Br	2	-11	-10	0	41
$\delta I = \delta RI_{PhCH2CH2-x} - SI$					
OH	17	35	28	26	36
CN	0	-1	-5	2	1
CO_2R	14	16	15	19	20
Cl	-14	-13	-8	-10	-14
Br	-5	-10	-10	-0	-15

Table 4.21: Interaction increments for interaction of aliphatic substituents with the aromatic ring in acetonitrile eluents

Substituent	Interaction increment Acetonitrile proportion (%) 30	40	50	60	70	80
$\delta I = \delta RI_{PhCH2-x} - SI$						
$CONH_2$	19	27	30	31	49	47
OH	61	72	80	82	96	97
CN	65	70	73	77	85	86
CO_2R	29	32	35	34	36	33
COR	3	9	12	14	13	6
Cl	19	21	20	16	17	11
Br	12	11	9	4	1	-8
$\delta I = \delta RI_{PhCH2CH2-x} - SI$						
OH	27	33	37	37	48	46
CN	17	18	18	18	22	18
CO_2R	23	23	25	25	27	23
Cl	-2	-2	-4	-10	-10	-20
Br	-3	-4	-5	-12	-16	-26

single length of alkyl chain it was not possible to derive their interaction increments. As with the calculation of substituent indices the interaction indices will be calculated from the regression equations relating the change in interaction increment to eluent composition (Tables 4.22 and 4.23). If the interaction increment was less than 10 it was regarded as being not significant and if the change across the eluent range was less than 10 a mean value was used rather than a fitted regression equation.

The largest interaction increments (δI +55 to +97) were observed for the nitrile group on the first carbon where there was a large increase in retention relative to the substituent index i.e. the nitrile group has a considerably less negative effect on the retention index than would be expected. In methanol the effect was very short range with the interaction increment being insignificant at $C_{n\ominus} > 1$ (Table 4.20). However, in acetonitrile the interaction increments were significant for both positions of the substituent (Table 4.21) but there was a rapid decrease in the interaction increment as the distance from the ring increased. The interaction increment for the substituent on ethylbenzene was approximately 25% that of the first.

A similar reduction in interaction increment was also observed with the hydroxyl group although in this case the reduction was less rapid with the value at $C_{n\ominus} = 2$ being about 50% of the value that for $C_{n\ominus} = 1$ in both methanol and acetonitrile. Unlike the nitrile substituent a significant interaction increment was observed in both methanol and acetonitrile for $C_{n\ominus} = 2$.

The ester group differed from the previous substituents in that the difference in the interaction increment between

Table 4.22: Regression equations relating interaction increments to methanol concentration

$II = ax^2 + bx + c$ x = % modifier

Substituent	Coefficients of interaction index equations a	b	c
II_{PhCH_2-x}			
$CONH_2$	0.0186	-1.809	56
OH	-0.0007	0.996	-3
CN	0.0221	-2.327	113
CO_2R	0	0	0
COR	0	0	-12
Cl	-0.0143	1.414	-31
Br	0.0836	-9.139	235
$II_{PhCH_2CH_2-x}$			
OH	-0.0079	1.233	-16
CN	0	0	0
CO_2R	0	0	17
Cl	0	0	-12
Br	-0.0071	0.757	-26

Table 4.23: Regression equations relating interaction increments to acetonitrile concentration

$II = ax^2 + bx + c$ x = % modifier

Substituent	Coefficients of interaction index equations a	b	c
II_{PhCH_2-x}			
$CONH_2$	0.0018	0.395	6
OH	-0.0046	1.236	29
CN	0	0.440	52
CO_2R	0	0	33
COR	0	0	10
Cl	0	0	17
Br	-0.0079	0.479	5
$II_{PhCH_2CH_2-x}$			
OH	-0.0021	0.636	10
CN	0	0	19
CO_2R	0	0	24
Cl	-0.0075	0.482	-10
Br	-0.0101	0.668	-14

the two lengths of alkyl chain was only very small. In methanol the interaction increment was actually larger for the substituent further from the ring, contrary to the expected sequence. When compared to the hydroxyl and nitrile group the interaction increments were considerably smaller and so any longer range effects would probably not have a significant effect on the retention index. However, in acetonitrile the interaction increments were sufficiently large to be significant.

Smaller effects were observed with the amide group, however, as the longest chain examined contained only 2 carbon atoms this may be at least partially due to the interaction increment, which would be expected for the second carbon.

The halides (Br and Cl) behaved very similarly to each other. In methanol and acetonitrile the interaction increments were less than those discussed previously and were probably not significant beyond the first carbon.

The final group for which it was possible to derive interaction increments was the aliphatic ketone group (COR), in acetonitrile there were only very small differences between the second and first carbon group. This functional group would be expected to behave in a manner similar to the ester functional group and the interaction increments derived were very similar to the difference between the interaction increments for the ester on the first and second carbons. At this time insufficient information was available to conclude that, at least in acetonitrile the same interaction index could be used for both functional groups. In methanol the situation was very different, the interaction increment was considerably larger than the ester

term and also negative. This was clearly very different to the previously observed values and suggests that a much larger interaction or combinations of interactions was occurring in the methanol eluents with low organic modifier content.

Although general comments on the size of the interactions have been made, it was difficult to correlate these with any properties of the substituents. Few substituent parameters attempt to account for differences which can occur in compounds containing both aromatic and aliphatic groups. In extending the concept of calculating log *P* by summation of π Leo[92] recently suggested a method to account for the various interactions. Interactions due to multiple substitution on a ring will be discussed in a later chapter but it was noted that a term , F_{α}, was required to account for the difference between aliphatic compounds and aromatic compounds, and compounds containing both aromatic and aliphatic groups. It was suggested that the upper limit of this term was at $C_{n\sigma}$= 3 and this agrees with the assumptions made in this chapter. No attempt was however made to quantify the F_{α} in terms of the substituents but a single overall value was applied to all compounds. It would appear that this approach was not valid as in this work considerable variation in the interaction increments was observed for different substituents.

Although further study might lead to the possibility of quantifying these interaction increments using non-empirical properties of the substituents or at least generalised interaction terms for different classes of compounds, this was not possible with this current set of data.

4.7 USE OF THE SUBSTITUENT INDEX EQUATIONS AND INTERACTION INDEX EQUATIONS

The substituent index equations and the interaction index equations described in this chapter will be used in the prediction system to predict the retention indices of compounds.

CHAPTER 5

INVESTIGATION OF THE RETENTION BEHAVIOUR OF ALIPHATIC COMPOUNDS WITH ISOMERIC SUBSTITUENTS AND DISUBSTITUTED ALIPHATIC COMPOUNDS

The retention behaviour isomeric alkylbenzenes and phenylpropanols were examined to determine whether chain branching had an effect on the retention index and whether the hydroxyl substituent index was the same for primary, secondary and tertiary isomers. Two disubstituted alkylbenzenes were also studied to determine the importance of substituent interactions in disubstituted aliphatic compounds.

5.1 EXAMINATION OF THE EFFECT OF CHAIN BRANCHING ON THE RETENTION OF ALKYLBENZENES

The capacity factors (Table 5.1) and retention indices (Table 5.2) of six isomeric alkylbenzenes (listed below) were determined over the eluent ranges 40 - 80% methanol and 30 - 80% acetonitrile. In 40% methanol the isomeric butylbenzenes were omitted due to the excessive retention times.

n-propylbenzene

$PhCH_2CH_2CH_3$

i-propylbenzene

$PhCH(CH_3)CH_3$

n-butylbenzene

$PhCH_2CH_2CH_2CH_3$

i-butylbenzene

$PhCH_2CH(CH_3)CH_3$

t-butylbenzene

$PhC(CH_3)_3$.

s-butylbenzene

$PhCH(CH_3)CH_2CH_3$

Table 5.1: Capacity factors of isomeric alkylbenzenes

Compound	Capacity factor Organic modifier proportion (%) 30	40	50	60	70	80
methanol-buffer						
ethylbenzene		65.28	27.28	13.73	5.02	2.35
i-propylbenzene		130.97	51.64	19.16	7.41	3.08
n-propylbenzene		165.50	59.57	24.60	8.19	3.38
t-butylbenzene		-	76.08	27.61	9.69	3.81
s-butylbenzene		-	95.77	32.85	11.16	4.38
i-butylbenzene		-	123.12	38.89	12.84	4.72
n-butylbenzene		-	130.97	41.20	13.37	4.95
acetonitrile - buffer						
ethylbenzene	59.53	22.38	9.46	4.66	2.67	1.58
i-propylbenzene	131.26	35.92	14.36	6.70	3.50	2.06
n-propylbenzene	137.23	42.42	15.83	7.17	3.88	2.15
t-butylbenzene	216.36	53.09	19.57	8.77	4.07	2.50
s-butylbenzene	290.89	65.43	23.39	10.28	4.69	2.82
i-butylbenzene	334.41	73.41	25.71	11.01	5.38	3.02
n-butylbenzene	308.86	79.97	27.39	11.81	5.63	2.94

Table 5.2: Retention indices of isomeric alkylbenzenes

Compound	Retention index Organic modifier proportion (%) 30	40	50	60	70	80
methanol-buffer						
ethylbenzene		1075	1100	1126	1151	1177
i-propylbenzene		1154	1179	1202	1224	1246
n-propylbenzene		1182	1204	1227	1256	1282
t-butylbenzene		-	1251	1269	1290	1313
s-butylbenzene		-	1280	1298	1320	1341
i-butylbenzene		-	1296	1317	1341	1362
n-butylbenzene		-	1317	1334	1358	1384
acetonitrile - buffer						
ethylbenzene	1095	1110	1121	1130	1137	1144
i-propylbenzene	1169	1186	1196	1203	1211	1215
n-propylbenzene	1200	1212	1223	1233	1244	1253
t-butylbenzene	1230	1248	1256	1264	1272	1280
s-butylbenzene	1266	1281	1291	1302	1312	1325
i-butylbenzene	1283	1299	1310	1322	1333	1344
n-butylbenzene	1302	1313	1322	1335	1346	1361

If chain branching had no effect on retention then the retention index of the branched chain isomers would be expected to be the same as the straight chain isomer but in each case a wide range of retention index values were found. The branched compounds were eluted more rapidly than the isomeric n-alkyl chains and the elution time decreased as the degree of branching increased. Small differences might be expected to arise with i-propylbenzene and s- and t-butylbenzene, where a methyl group has been substituted on the carbon adjacent to the ring. Studies on the normal alkylbenzenes suggested that replacement of a hydrogen on the benzylic methyl group with a methyl group contributed a smaller retention increment than subsequent additions of methylene groups (Section 4.3).

5.1.1 Substituent Contributions of the Methyl/Methylene Groups in Isomeric Alkylbenzenes

To examine the effects of the addition of a secondary methyl group or a tertiary methyl group to the alkyl chain, a parent alkylbenzene has been defined, in each case based on the longest possible straight chain. Ethylbenzene has been used as the parent for i-propylbenzene and t-butylbenzene and n-propylbenzene has been used for i-butylbenzene and s-butylbenzene. The retention index increments (δRI) (Table 5.3) were calculated by subtracting the calculated retention index (Tables 4.10 and 4.11) of the parent alkylbenzene from the corresponding branched alkylbenzene (from Table 5.2).

The increments for the secondary and tertiary methyl groups substituted on the benzylic carbon (C-1 in

Table 5.3: Retention index increments for methyl substituents on alkyl chains

Parent	Alkylbenzene	Carbon Substituted	Retention increment Organic modifier proportion (%)					
			30	40	50	60	70	80
methanol-buffer								
ethylbenzene								
	i-propylbenzene	1	-	81	78	76	75	76
	t-butylbenzene	1	-	-	150	143	141	143
	(hence for each methyl		-	-	75	72	71	72)
n-propylbenzene								
	s-butylbenzene	1	-	-	79	72	71	71
	i-butylbenzene	2	-	-	95	91	92	92
acetonitrile-buffer								
ethylbenzene								
	i-propylbenzene	1	71	71	68	64	65	64
	t-butylbenzene	1	132	133	128	122	124	129
	(hence for each methyl		66	67	64	61	62	65)
n-propylbenzene								
	s-butylbenzene	1	68	66	63	63	66	74
	i-butylbenzene	2	85	84	82	83	87	93

Table 5.4: Interaction increments calculated for secondary and tertiary methyl groups substituted on the benzylic carbon or subsequent carbon calculated taking into account the interaction index for the second carbon of an alkyl chain

Carbon chain	Interaction increment Organic modifier proportion (%)					
	30	40	50	60	70	80
methanol-buffer						
$CH(CH_3)_2$	-	-7	-10	-12	-13	-12
$CH(CH_3)_3$	-	-	-26	-33	-35	-33
(for 1 group	-	-	-13	-17	-18	-17)
$CH(CH_3)CH_2CH_3$	-	-	-9	-16	-17	-17
$CH_2CH(CH_3)_2$	-	-	-5	-9	-8	-8
mean	-	-7	-9	-14	-14	-14
acetonitrile- buffer						
$CH(CH_3)_2$	-17	-17	-20	-24	-23	-24
$CH(CH_3)_3$	-44	-43	-48	-54	-52	-47
(for 1 group	-22	-22	-24	-27	-26	-24)
$CH(CH_3)CH_2CH_3$	-20	-22	-25	-25	-22	-14
$CH_2CH(CH_3)_2$	-15	-16	-18	-17	-13	-7
mean	-19	-19	-22	-23	-21	-17

i-propylbenzene, s-butylbenzene and t-butylbenzene) were very similar (71 to 81 in methanol and 61 to 74 in acetonitrile). This suggested that the relative retentions of the isomers were not determined purely on steric grounds as the effect each of the two tertiary methyl groups in t-butylbenzene might be expected to be greater than the secondary methyl group. The other superimposed effects may be due to hyperconjugation of the alkyl groups with the ring, where the electron releasing effects of the tertiary methyl group would be greater than the equivalent primary group. The polarity of these compounds would therefore be expected to be larger than the corresponding secondary and primary carbons and t-butylbenzene had the smallest retention index of the three isomers. The retention index increment of the methyl group substituted onto the second carbon (91 to 95 in methanol and 82 to 93 in acetonitrile) was always larger than the effect of substitution on the benzylic carbon but was still less than the defined methylene increment of 100.

Previously in studying the retention increments for n-alkyl chains it had been found that substituting a methyl group onto the benzylic carbon caused a reduction in the retention index (as compared to the defined value of 100) of -12. The interaction increments to account for the difference between primary methyl (SI = 100) and secondary and tertiary methyl groups have therefore been calculated allowing for this interaction index (II_{PhCH-R}) for each methyl/methylene group substituted on the benzylic carbon and assuming a nominal value of 100 for each group (e.g. Equation 5.1 for i-propylbenzene or s-butylbenzene and 5.2 for t-butylbenzene).

$$\delta I = \delta RI - (SI_{CH3} + II_{PhCH-R}) \qquad 5.1$$

$$\delta I = \delta RI - (2SI_{CH3} + 2II_{PhCH-R}) \qquad 5.2$$

The values obtained for the different compounds (Table 5.4) were similar and the differences are probably not significant. The mean values, at each eluent composition, have therefore been used to calculate the interaction indices for a non-primary methyl groups. As the change across the eluent ranges was less than 10 units (-7 to -14 in methanol and -17 to -23 in acetonitrile) the mean values have been used rather than fitted quadratic equations (Table 5.5). These corrections have been subsequently applied to compensate for the effect of branching an alkyl side chains.

5.1.2 Relationship between retention increment and structural parameters

The octanol-water substituent constants (π) have been reported for most of the isomeric alkylbenzene chains (Table 5.6). In all cases the π value for the straight chain isomer was higher than for the branched chains which agrees with the sizes of the retention increments. There are insufficient values available to determine whether the π values could predict the order of elution of the branched isomers. However, the values available for t-butyl and s-butyl would correctly predict the relative size of the retention increments of these alkyl chains.

It is also possible to calculate the molecular connectivity indices for the alkyl chains as described by

Kier[17] (Table 5.6). The molecular connectivity indices for the straight chain isomers were again higher than the corresponding branched chains and would correctly predict a larger retention increment for these chains. However, the relative order of elution of the sec and iso branched isomers would not have been predicted from the molecular connectivity indices. This agreed with previous findings of Smith[24] in which molecular connectivity indices could not predict the elution order of isomeric alkylbenzenes on an ODS-Hypersil column.

5.2 EXAMINATION OF RETENTION OF ISOMERIC PHENYLPROPANOLS

Five isomeric phenylpropanols (listed below) were examined to discover whether the contribution of the non-primary hydroxyl groups to the retention index was the same as a primary hydroxyl and also whether the effects of alkyl chain branching observed with the alkylbenzenes applied to these compounds.

3-phenyl-1-propanol	$PhCH_2CH_2CH_2OH$
2-phenyl-2-propanol	$PhC(CH_3)_2OH$
2-phenyl-1-propanol	$PhCH(CH_3)CH_2OH$
1-phenyl-2-propanol	$PhCH_2CH(OH)CH_3$
1-phenyl-1-propanol	$PhCH(OH)CH_2CH_3$

Table 5.5: Interaction index equations for secondary and tertiary methyl groups

$$II_{branch} = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Modifier	Coefficients of regression equations		
	a	b	c
methanol	0	0	-12
acetonitrile	0	0	-20

Table 5.6: Structural parameters of isomeric alkylbenzenes

Chain	π[48]	$^1\chi$[17]
$-CH_3$	0.56	0.410
$-CH_2CH_3$	1.05	0.971
$-CH_2CH_2CH_3$	1.55	1.411
$-CH(CH_3)_2$	1.53	1.354
$-CH_2CH_2CH_2CH_3$	2.13	1.971
$-CH(CH_3)CH_2CH_3$	2.04	1.892
$-CH_2CH(CH_3)_2$	-	1.827
$-C(CH_3)_3$	1.98	1.661

The different compounds can be divided into groups depending on their structure, 3-phenyl-1-propanol, 1-phenyl-2-propanol and 1-phenyl-1-propanol have straight alkyl chains with hydroxyl groups substituted on different carbon atoms. 2-phenyl-1-propanol has a branched alkyl chain with a primary hydroxyl substituent as a terminal substituent on a carbon atom. The final compound, 2-phenyl-2-propanol, has a branched chain with a tertiary hydroxyl group.

The capacity factors of these compounds (Table 5.7) have been used to calculate retention indices (Table 5.8). There were differences between the retention indices of the different isomers. Unlike the alkylbenzenes the isomeric compounds did not all elute earlier than the straight chain isomer (3-phenyl-1-propanol) which was perhaps not unexpected. In the previous chapter (Chapter 4) it has been seen that primary hydroxyl groups closer to the ring had a smaller effect on retention index than those further from the ring. It appeared that a similar effect was also true for the secondary hydroxyl groups. The order of elution was also different in the two eluents probably due to selectivity differences.

5.2.1 Retention increments calculated for hydroxyl groups

The hydroxyl retention index increments have been calculated from the retention indices using the Equations 5.3 to 5.6. In these equations it has been assumed that the interactions of hydroxyl groups with the aromatic ring (II_{1-OH} and II_{2-OH}, from Table 4.23 and 4.24) were not dependent on the "type" of hydroxyl group but only on the distance from the ring (i.e C-1 and C-2 from phenyl). When

Table 5.7: Capacity factors of isomeric phenylpropanols

Compound	Capacity factor Organic modifier proportion (%)					
	30	40	50	60	70	80
methanol - buffer						
2-phenyl-2-propanol	-	7.83	3.31	1.70	1.01	0.54
2-phenyl-1-propanol	-	8.15	3.48	1.72	1.02	0.55
1-phenyl-2-propanol	-	8.18	3.52	1.75	1.02	0.57
3-phenyl-1-propanol	-	9.12	3.75	1.78	1.03	0.56
1-phenyl-1-propanol	-	9.81	4.04	1.95	1.11	0.60
acetonitrile - buffer						
1-phenyl-2-propanol	4.61	2.23	1.41	0.92	0.64	0.43
2-phenyl-2-propanol	4.73	2.29	1.48	0.97	0.66	0.42
2-phenyl-1-propanol	4.96	2.35	1.45	0.94	0.63	0.42
3-phenyl-1-propanol	5.18	2.37	1.42	0.91	0.62	0.41
1-phenyl-1-propanol	6.00	2.79	1.74	1.10	0.73	0.47

Table 5.8: Retention indices of isomeric phenylpropanols

Compound	Retention index Organic modifier proportion (%)					
	30	40	50	60	70	80
methanol - buffer						
3-phenyl-1-propanol	-	840	835	828	813	793
2-phenyl-2-propanol	-	825	823	819	809	789
2-phenyl-1-propanol	-	830	827	821	812	792
1-phenyl-2-propanol	-	830	829	824	812	797
1-phenyl-1-propanol	-	851	847	842	830	812
acetonitrile - buffer						
3-phenyl-1-propanol	782	756	739	728	731	735
1-phenyl-2-propanol	769	755	740	734	734	748
2-phenyl-2-propanol	772	760	750	745	742	746
2-phenyl-1-propanol	778	763	746	738	733	740
1-phenyl-1-propanol	801	791	780	775	773	777

the alkyl chain was branched (2-phenyl-2-propanol and 2-phenyl-1-propanol) the interaction index for branched alkyl chains ($II_{\beta-CH_3}$) has been applied as well as the interaction index for the substitution of methyl groups on the benzylic carbon (I_{PhCH-R}) if appropriate.

For the compound 2-phenyl-1-propanol which has a primary hydroxyl group and a branched chain the hydroxyl retention index increment was calculated from equation 5.3:-

$$\delta RI = RI - (PI + 3SI_{CH_2} + 2II_{PhCH_2-R} + II_{2-OH} + II_{\beta-CH_3}) \qquad 5.3$$

The retention increments (Table 5.9) were very similar to those obtained previously for primary hydroxyl groups on straight chains and the values for 3-phenyl-1-propanol included for comparison (Tables 4.10 and 4.11).

The two compounds 1-phenyl-2-propanol and 1-phenyl-1-propanol both contain straight alkyl chains with secondary hydroxyl groups substituted on different carbon atoms. The retention increments have been calculated from equation 5.4 for 1-phenyl-2-propanol and equation 5.5 for 1-phenyl-1-propanol.

$$\delta RI = RI - (PI + 3SI_{CH_2} + II_{PhCH_2-R} + II_{2-OH}) \qquad 5.4$$

$$\delta RI = RI - (PI + 3SI_{CH_2} + II_{PhCH_2-R} + II_{1-OH}) \qquad 5.5$$

The final phenylpropanol, 2-phenyl-2-propanol, contains a branched chain and a tertiary hydroxyl group. The retention increment can be calculated from equation 5.6.

$$\delta RI = RI - (PI + 3SI_{CH_2} + 2II_{PhCH_2-R} + II_{1-OH} + 2II_{3-CH_3}) \quad 5.6$$

The calculated retention index increments for the two secondary and the tertiary hydroxyl groups were very similar but differed significantly from those for the primary hydroxyl group in 2-phenyl-1-propanol. The similarity of the results for these compounds confirms the assumption that the use of the interaction terms for the phenyl group is valid. A single substituent index (in each organic modifier) has therefore been calculated from the mean retention increments for the secondary and tertiary hydroxyl groups (Table 5.10) and will be used in the prediction of non-primary hydroxyl groups.

5.3 RETENTION BEHAVIOUR OF SELECTED COMPOUNDS WITH TWO SUBSTITUENTS ON THE ALKYL CHAIN

Two compounds (listed below) were examined in which the adjacent aliphatic substituents might be expected to undergo hydrogen bonding to test the extent of the effect of the interactions. However, the number of compounds studied was not sufficient to allow any conclusions about interactions to be made and the results have not been incorporated into the retention prediction system at this stage.

1-phenyl-1,2-ethanediol $PhCH(OH)CH_2(OH)$

ethyl DL-mandelate $PhCH(OH)CO_2C_2H_5$.

Table 5.9: Retention index increments calculated for hydroxyl groups

Compound	Retention index increment Organic modifier proportion (%) 30	40	50	60	70	80
methanol - buffer						
3-phenyl-1-propanol		-335	-366	-391	-436	-477
2-phenyl-1-propanol		-342	-378	-413	-447	-500
1-phenyl-1-propanol		-358	-399	-438	-482	-530
1-phenyl-2-propanol		-364	-398	-432	-469	-517
2-phenyl-2-propanol		-362	-401	-439	-451	-531
$\delta RI_{\bullet-OH}$ (mean)		-361	-399	-436	-467	-526
acetonitrile - buffer						
3-phenyl-1-propanol	-416	-459	-489	-511	-515	-516
2-phenyl-1-propanol	-417	-454	-489	-512	-527	-528
1-phenyl-1-propanol	-459	-503	-500	-551	-556	-572
1-phenyl-2-propanol	-456	-492	-514	-546	-556	-550
2-phenyl-2-propanol	-458	-496	-527	-551	-567	-573
$\delta RI_{\bullet-OH}$ (mean)	-458	-497	-514	-549	-560	-565

Table 5.10: Substituent index equations for secondary and tertiary hydroxyl groups

Calculations from the mean values in Table 5.9.

$$SI_{\bullet\bullet c-OH} = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Modifier	Coefficients of regression equation a	b	c
methanol	-0.0257	-0.894	-286
acetonitrile	0.0346	-5.979	-310

The capacity factors of these compounds (Table 5.11) have been used to calculate the retention indices (Table 5.12). The anticipated retention indices of these compounds, if there was no hydrogen bonding effect, can be calculated as the sum of the contributions of the different parts using the equations given below.

1-phenyl-1,2-ethanediol

$$RI_{calc} = PI + 2SI_{R} + SI_{prim-OH} + SI_{sec-OH} + II_{PhCH2-R} + II_{1-OH} + II_{2-OH} \qquad 5.10$$

ethyl DL-mandelate

$$RI_{calc} = PI + 3SI_{R} + SI_{sec-OH} + SI_{A1-CO2R} + II_{1-OH} + II_{1-CO2R} \qquad 5.11$$

The difference between calculated retention indices obtained using these equations and the experimental retention indices, the interaction increments (Table 5.13), were a measure of the size of the interactions. These hydrogen bonding effects differed between the two sets of substituents and were also dependent on the organic modifier and organic modifier concentration. It may be possible to generalise these results to other similar pairings for example other carbonyl-hydroxyl pairings, or cases in which the hydroxyl is replaced by a amino group. The values are comparable to those determined for aromatic hydrogen bonding (100 to 300 units, Chapter 7.3.2).

However insufficient compounds have been examined to apply these interactions in the prediction system and a

Table 5.11: Capacity factors of 1-phenyl-1,2-ethanediol and ethyl DL-mandelate

Compound	Capacity factor Organic modifier proportion (%) 30	40	50	60	70	80
methanol - buffer						
1-phenyl-1,2-ethanediol		1.20	0.72	0.48	0.35	0.23
ethyl DL-mandelate		5.45	2.30	1.15	0.69	0.41
acetonitrile - buffer						
1-phenyl-1,2-ethanediol	0.77	0.54	0.41	0.32	0.26	0.25
ethyl DL-mandelate	4.33	1.98	1.38	0.83	0.56	0.36

Table 5.12: Retention indices of disubstituted compounds

Compound	Retention index Organic modifier proportion (%) 30	40	50	60	70	80
methanol-buffer						
1-phenyl-1,2-ethanediol	---	620	612	601	583	546
1-phenyl-1,2-ethanediol*	---	431	416	388	347	293
ethyl DL mandelate	---	792	773	754	728	696
ethyl DL mandelate*	---	669	643	613	574	531
acetonitrile-buffer						
1-phenyl 1,2 ethanediol	548	531	496	481	480	569
1-phenyl-1,2-ethanediol*	333	280	243	224	222	237
ethyl DL-mandelate	761	741	723	710	696	694
ethyl DL-mandelate*	628	592	561	535	529	518

* value calculated as the sum of *PI* and *SI*

Table 5.13: Interaction increments for disubstituted aliphatic compounds

Substituent Pairs	Interaction increment Organic modifier proportion (%) 30	40	50	60	70	80
OH-OH (MeOH:Buffer)	---	158	157	182	221	232
OH-OH (MeCN:Buffer)	191	223	224	232	218	298
OH-CO_2CH_3 (MeOH:Buffer)	---	123	130	141	154	165
OH-CO_2CH_3 (MeCN:Buffer)	133	149	162	175	167	176

larger number of disubstituted compounds would have to be examined to obtain meaningful interaction indices. Consequently interaction index equations have not been calculated or included in the the database as insufficient information was available to ensure that they were generally reliable.

CHAPTER 6

SUBSTITUENT INDICES OF AROMATIC SUBSTITUENTS

In this chapter the determination of substituent indices of aromatic substituents will be described. The indices have been calculated using 13 mono-substituted benzenes to give values in the absence of interactions with other substituents. The model compounds contained the same functional series of groups as have been examined in the study of aliphatic substituents with the addition of the nitro and groups. The substituent index equations derived in this chapter have then been used to calculate the interaction indices for disubstituted benzenes (Chapter 7).

6.1 RETENTION INDICES OF MONO-SUBSTITUTED MODEL COMPOUNDS

The capacity factors of the model compounds (Table 6.1 measured over the eluent ranges described previously (40 - 90% methanol and 30 - 90% acetonitrile) were used to calculate retention indices (Table 6.2). Acetophenone, which is one of the retention index standards, has also been included as a model compound to derive a substituent index for the aromatic ketone group. As with the aliphatic substituted model compounds the changes in retention indices were generally not linear but appeared to be systematic between 30 and 80% acetonitrile and 40 and 80% methanol (Figure 6.1 and 6.2). However with a few model compounds (e.g. aniline, phenol, benzamide) there were relatively large changes in the retention index in the eluents

Table 6.1: Capacity factors of monofunctional aromatic test compounds in eluents containing different proportions of methanol and acetonitrile

Compound	Capacity factor (*k'*) Organic modifier proportion (%) 30	40	50	60	70	80	90
methanol-buffer							
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
benzonitrile		5.01	2.83	1.37	0.86	0.50	0.35
biphenyl		204.00	62.30	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
acetonitrile-buffer							
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
benzonitrile	5.86	3.27	1.82	1.15	0.74	0.55	0.28
biphenyl	154.50	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

Table 6.2: Retention indices of model mono-substituted aromatic compounds in methanol and acetonitrile eluents

Compound	Retention index Organic modifier proportion (%) 30	40	50	60	70	80	90
methanol-buffer							
parent index (benzene)[a]	-	885	913	938	961	982	1000
acetophenone[b]	-	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
benzonitrile	-	776	788	775	774	760	736
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	-	680	683	680	671	650	582
toluene	-	987	1019	1039	1065	1095	1132
acetonitrile-buffer							
parent index (benzene)[a]	910	927	940	951	958	963	966
acetophenone[b]	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
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] From Table 4.2
[b] calculated for k' taken from alkylarylketone test mixture, nominal value 800

Figure 6.1: Change in retention indices of a selection of aromatic compounds with proportion of methanol

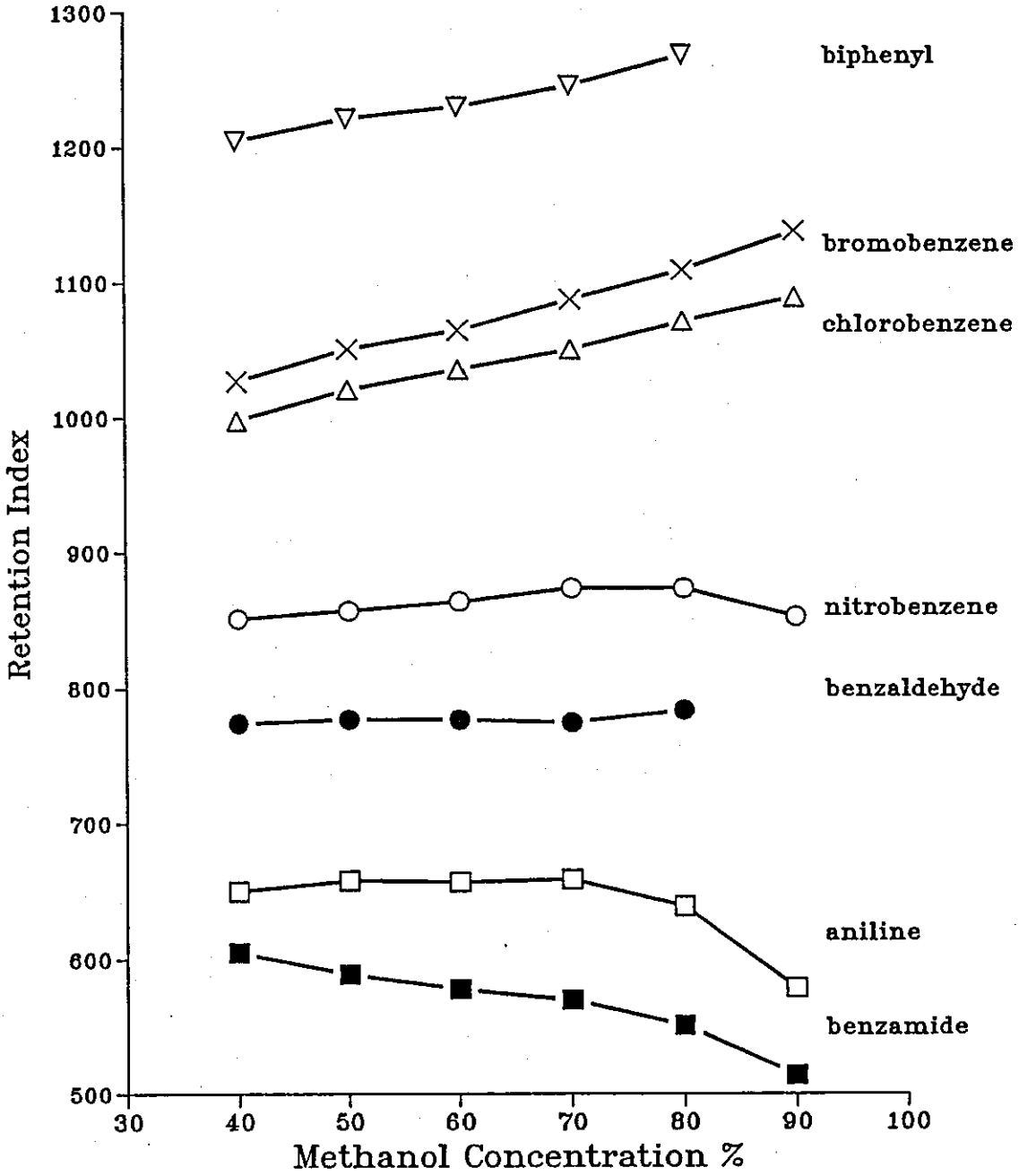
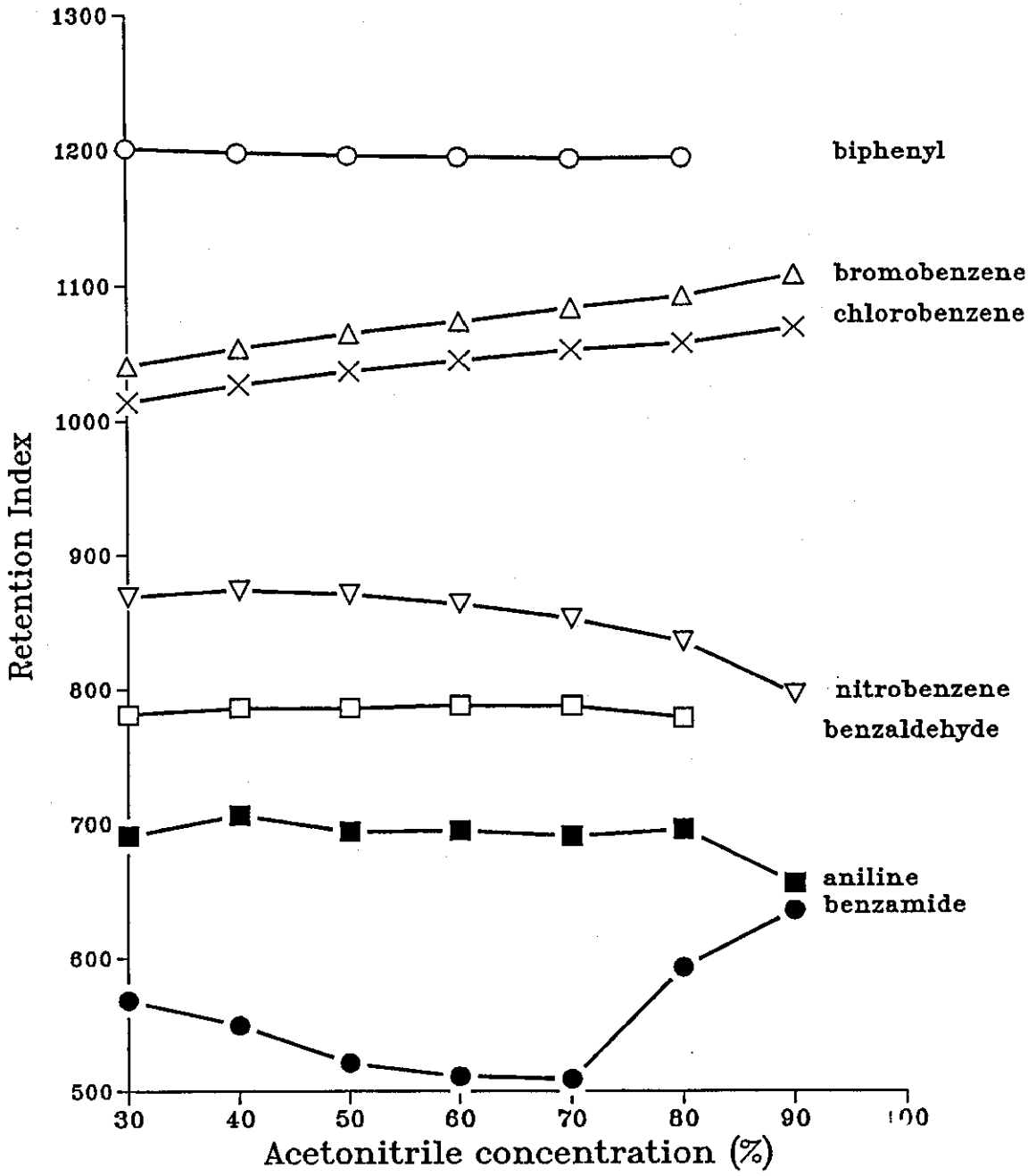


Figure 6.2: Change in retention indices of a selection of aromatic compounds with proportion of acetonitrile



containing 90% organic modifier. Because of the sudden change and the uncertainty in the capacity factor at 90% organic modifier these results were excluded from subsequent calculations and will be disregarded in the discussion.

The aromatic amine (aniline) did not show any of the "odd" behaviour patterns observed with the more basic aliphatic amines (Chapter 4, Section 4.2). Aniline is weakly basic and has a pKa of 4.63[165] and would therefore not be expected to be ionised at the pH of the mobile phase and should not be retained by an ion exchange mechanism.

In previous studies on the use of retention indices on the alkylarylketone scale Smith and co-workers[134,145,146] found that the retention index of methyl benzoate was virtually unchanged across the eluent ranges and was independent of the column packing material. This was consistent with this compound belonging to the same Snyder[171] selectivity group as the alkylarylketones. Although the values reported here differ slightly from those previously reported by Smith[134,145] on a Hypersil-ODS column (*RI* = 916-963 in methanol[134] and *RI* = 880-903 in acetonitrile[145]) the change in retention index across the eluent range up to 80% modifier was only small (*RI* = 899 to 914 in methanol - buffer and *RI* = 890 to 897 in acetonitrile - buffer).

Benzaldehyde might be expected to behave as the zeroth member of the alkylarylketone homologous series and therefore have a retention index approximately 100 less than that defined for acetophenone. The experimental retention index was found to be in the range *RI*= 774 to 784 in methanol and *RI* = 779 to 788 in acetonitrile containing eluents. These values were similar to those found previously

by Smith[134] in eluents containing methanol (*RI* = 758 to 776 on a Hypersil-ODS column).

6.2 SUBSTITUENT INDICES FOR SINGLE AROMATIC SUBSTITUENTS

Retention index increments (δRI) for the substituents were earlier defined as the difference between the retention index of the substituted compound and the parent index. For the aromatic substituents they were therefore calculated using the equation.

$$\delta RI = RI_{PhX} - PI$$

Where *PI* is the parent index calculated from the regression equation described earlier (Table 6.2). Between 30 and 80 % acetonitrile and 40 and 80 % methanol the changes in the retention index increments for all the functional group were systematic although not linear (Tables 6.3 and 6.4).

The experimental values for the methyl group , from toluene, were close to the defined value of 100 and were usually within the experimental error of +/- 10 . The only exception was at 80% methanol, where there was an increased uncertainty in the calculated values and the defined value of 100 has therefore been used as the substituent index.

Among the substituents were several mixed alkyl-aryl functional groups ($COCH_3$, CO_2CH_3 and OCH_3) and as with the aliphatic functional groups in each case the methyl group was defined as having a substituent index of 100 and this was subtracted from the calculated retention index increment to obtain the values for the COR group etc.

Table 6.3: Retention index increments of aromatic functional groups in methanol eluents

Substituent	Retention index increment				
	Methanol proportion (%)				
	40	50	60	70	80
$CONH_2$	-280	-324	-360	-391	-431
NH_2	-235	-255	-281	-302	-343
OH	-205	-230	-258	-290	-332
CHO	-111	-136	-161	-186	-198
CN	-109	-133	-163	-187	-222
$COCH_3$[a]	-80	-107	-135	-158	-178
$COCH_3$[b]	-85	-113	-138	-162	-182
(COR[*]	-185	-213	-238	-262	-282)
NO_2	- 34	- 56	- 74	- 87	-108
OCH_3	-1	-9	-20	-27	-28
(OR[*]	-101	-109	-120	-127	-128)
CO_2CH_3	14	-9	-34	-51	-68
(CO_2R[*]	-86	-109	-134	-151	-168)
H	0	0	0	0	0
CH_3[a]	102	106	101	104	113
CH_3[b]	100	100	100	100	100
Cl	113	108	98	90	90
Br	142	138	127	127	128
Ph	320	309	293	286	288

[a] based on experimental values of retention indices
[b] based on defined value ($-COCH_3$ RI = 800, $-CH_3$ δRI = 100) and used for substituent index calculation
[*] calculated by subtracting 100 for methyl contribution

Table 6.4: Retention index increments for aromatic functional groups in acetonitrile eluents

Substituent	Retention index increment Acetonitrile proportion (%)					
	30	40	50	60	70	80
$CONH_2$	-342	-378	-419	-440	-449	-370
NH_2	-219	-221	-246	-256	-267	-267
OH	-215	-240	-266	-291	-319	-318
CHO	-129	-141	-154	-163	-170	-184
$COCH_3$[a]	-110	-129	-142	-153	-159	-160
$COCH_3$[b]	-110	-127	-140	-151	-158	-163
(COR[*]	-210	-227	-240	-251	-258	-263)
CN	-96	-110	-127	-143	-151	-175
NO_2	-41	-53	-69	-87	-105	-127
CO_2CH_3	-20	-37	-40	-57	-64	-66
(CO_2R[*]	-120	-137	-140	-157	-164	-166)
OCH_3	-10	-19	-28	-36	-41	-50
(OR[*]	-110	-119	-128	-136	-141	-150)
H	0	0	0	0	0	0
CH_3[a]	95	95	96	94	95	95
CH_3[b]	100	100	100	100	100	100
Cl	104	100	97	94	95	95
Br	131	127	125	123	126	130
Ph	291	271	256	244	236	232

[a] based on experimental values of retention indices
[b] based on defined value ($-COCH_3$ RI = 800, $-CH_3$ δRI = 100) and used for substituent index calculation
[*] calculated by subtracting 100 for methyl contribution

Table 6.5: Substituent index equations for aromatic substituents in methanol eluents (40 - 80 %)

$SI = ax^2 + bx + c$ x = % modifier

Substituent	Hansch[48] π	Coefficients of regression equation a	b	c
$CONH_2$	-1.49	0.0093	-4.804	-104
NH_2	-1.23	-0.0264	0.541	-215
OH	-0.67	-0.0271	0.117	-167
CHO	-0.65	0.0186	-4.469	39
CN	-0.57	-0.0114	-1.429	-34
COR	-0.55	0.0114	-3.791	-52
NO_2	-0.28	0.0050	-2.390	53
OR	-0.02	0.0129	-2.263	-30
CO_2R	-0.01	0.0143	-3.774	43
CH_3	0.56	0	0	100*
Cl	0.71	0.0086	-1.669	167
Br	0.86	0.0150	-2.190	207
Ph	1.96	0.0250	-3.870	436

* by definition

Table 6.6: Substituent index equations for aromatic substituents in acetonitrile eluents (30 - 80 %)

$SI = ax^2 + bx + c$ x = % modifier

Substituent	Hansch[48] π	Coefficients of regression equation a	b	c
$CONH_2$	-1.49	0.1260	-14.878	2
NH_2	-1.23	0.0118	-2.405	-153
OH	-0.67	0.0218	-4.616	-93
CHO	-0.65	0.0025	-1.335	-92
CN	-0.57	-0.0025	-1.251	-57
COR	-0.55	0.0150	-2.704	-143
NO_2	-0.28	-0.0104	-0.586	-14
OR	-0.02	0.0029	-1.097	-80
CO_2R	-0.01	0.0105	-2.096	-67
CH_3	0.56	0	0	100*
Cl	0.71	0	0	98
Br	0.86	0	0	127
Ph	1.96	0.0193	-3.299	372

* by definition

Although there was a small systematic error in the retention index of acetophenone compared to the nominal value the difference was less than the experimental uncertainty (+/- 10 units). Therefore the retention index increments and hence substituent indices have been calculated based on the defined value of *RI* = 800 rather than the experimental values.

Quadratic regression equations were fitted to the experimental data or defined values to calculate the coefficients of substituent index (*SI*) equations (Tables 6.5 and 6.6). For the bromide and chloride substituents the change in retention increment across the range of acetonitrile concentrations was less than the experimental uncertainty (Br δRI 131 - 123 and Cl δRI 104 - 94). For these two substituents the mean value was therefore used as the substituent index. The fitted curves were a good description of the experimental data for all the substituents as shown in Figures 6.3 and 6.4, for a selection of the functional groups, where the lines represent the calculated substituent indices and the points are the experimental retention index increments.

The coefficients of the substituent index equations have been written into a spreadsheet and will be used for retention prediction (see Chapter 8).

6.3 RELATIONSHIP BETWEEN SUBSTITUENT INDICES AND PHYSICO-CHEMICAL PROPERTIES

In a previous chapter (Chapter 4) there was found to be an approximately relationship between the substituent

Figure 6.3: Comparison of experimental substituent increments and calculated substituent indices in methanol

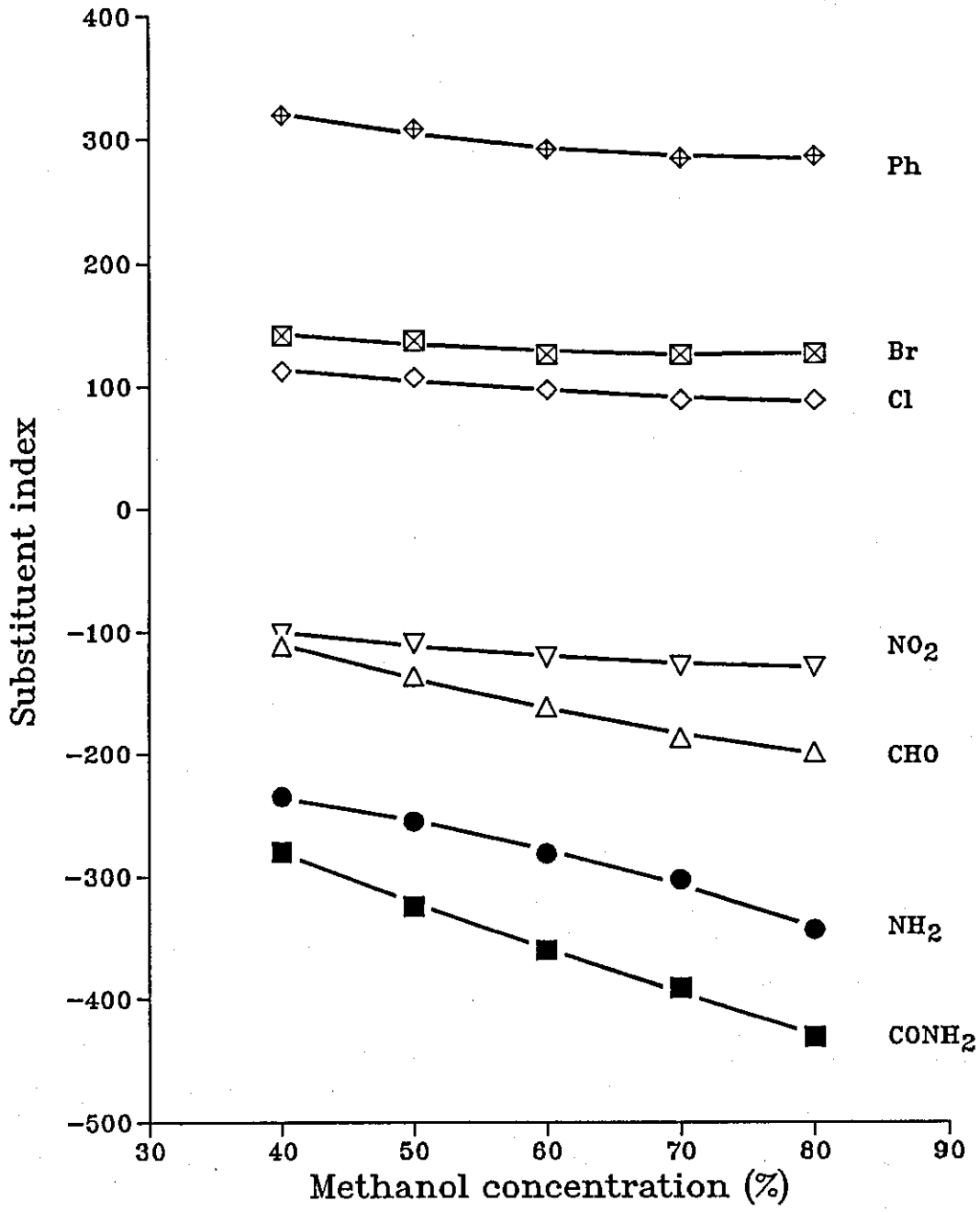
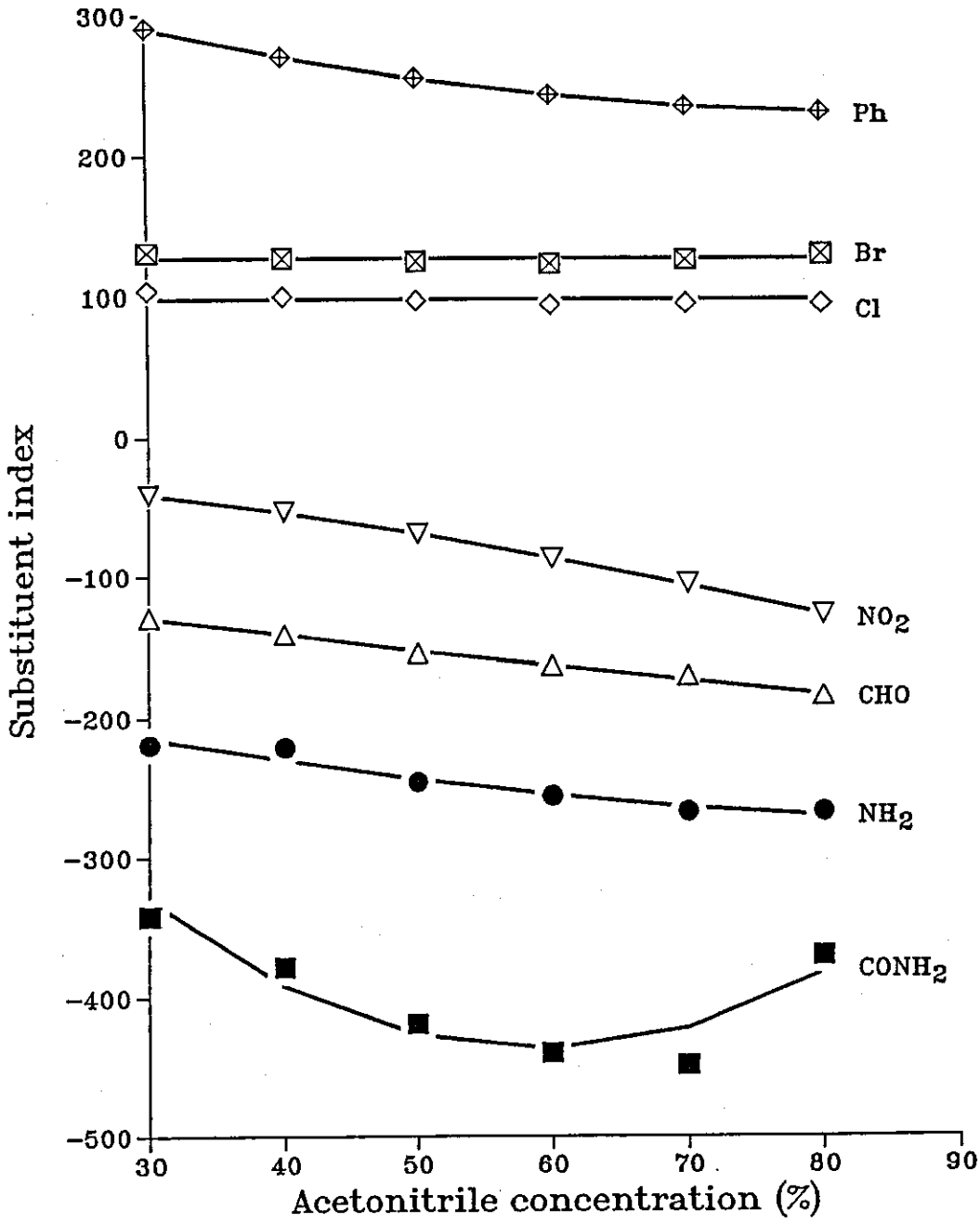


Figure 6.4: Comparison of experimental substituent increments and calculated substituent indices in acetonitrile



contribution to the octanol-water partition coefficients and the substituent indices for aliphatic substituents. In that chapter the relationship between the substituent index and solubility constants was also investigated. However, a full study of the relationship between aqueous solubility and *SI* cannot be carried out for the aromatic substituents because of a lack of reliable contributions, for the aromatic substituents, to aqueous solubility.

6.3.1 Relationship between *SI* and octanol-water partition coefficient

As was discussed earlier, within groups of closely related compounds log k' of analytes in reversed phase HPLC has frequently been linearly correlated with the octanol-water partition coefficient (log P) as both processes represent a liquid-liquid partition mechanism. The aliphatic substituent indices were found to be linearly related with the Hansch substituent contributions[48] to the octanol-water partition coefficient (π) (Chapter 4). A similar study was carried out with the aromatic substituents to determine whether the substituent indices of aromatic substituents varied linearly with π values. As π values were also available for benzylic functional groups (CH_2OH, CH_2CN, CH_2Cl and CH_2Br) these have been included in the correlation to investigate whether these groups could be treated as aromatic substituents. The retention increments for these compounds (Table 6.7) were calculated from the retention indices listed previously (Chapter 4, Table 4.6 and 4.7). These were used to obtain substituent index equations and substituent indices were used in the regression

calculations.

As with the aliphatic substituents a generally linear relationship was found between the π values (Tables 6.5, 6.6 and 6.7) and the substituent indices with correlation coefficients ranging from 0.970 to 0.985 in methanol and 0.960 to 0.968 in acetonitrile containing eluents (Table 6.8). An examination of the curve for *SI* vs π in methanol eluents (Figure 6.5) identified a single outlier the phenolic hydroxyl group. Differences in the relationship between log *P* and log *k'* between hydrogen bonding species and non-hydrogen bonding species have been noted previously[56]. The equivalent comparison of aromatic substituents in acetonitrile suggested two possible outliers, the phenolic hydroxyl and the amide again both are hydrogen bonding species (Figure 6.6).

In neither modifier system were the CH_2-X groupings outliers, suggesting that the π values for these groups were a good guide to the chemical properties. However, it is not intended to include their substituent indices in the database for aromatic substituents, rather they have been included as aliphatic substituents on an alkyl side chain and an interaction increment used to account for the proximity of the phenyl group (see Chapter 4).

The substituent π values could be used to estimate substituent indices of other substituents, however, as with the aliphatic substituents, there would be considerable uncertainty in the calculated values.

Table 6.7: Retention increments for benzyl substituents CH_2OH, CH_2CN, CH_2Cl and CH_2Br

Functional group	Hansch[48] π	Retention increment Organic modifier proportion (%) 30	40	50	60	70	80
methanol - buffer							
CH_2OH	-1.03	-	-196	-222	-240	-277	-307
CH_2CN	-0.57	-	-113	-147	-175	-233	-260
CH_2Cl	0.17	-	77	63	54	31	12
CH_2Br	0.79	-	106	91	81	69	77
acetonitrile - buffer							
CH_2OH	-1.03	-256	-287	-310	-327	-322	-318
CH_2CN	-0.57	-85	-110	-136	-162	-184	-212
CH_2Cl	0.17	89	72	53	32	15	-7
CH_2Br	0.79	116	98	80	60	44	23

Table 6.8: Regression coefficients for π vs substituent interaction index in methanol eluents

$$SI = a \times \pi + b$$

Modifier-buffer	Coefficients of regression equation a	b	Correlation coefficient
methanol-buffer			
40:60	-7.97	178	0.9849
50:50	-26.63	186	0.9839
60:40	-44.53	195	0.9816
70:30	-62.00	207	0.9789
80:20	-70.58	223	0.9607
acetonitrile-buffer			
30:70	-18.96	180	0.9739
40:60	-35.32	189	0.9661
50:50	-55.55	194	0.9573
60:40	-58.92	200	0.9588
70:30	-70.04	197	0.9622
80:20	-77.64	194	0.9627

Figure 6.5: Relationship between π values and substituent indices in methanol eluents

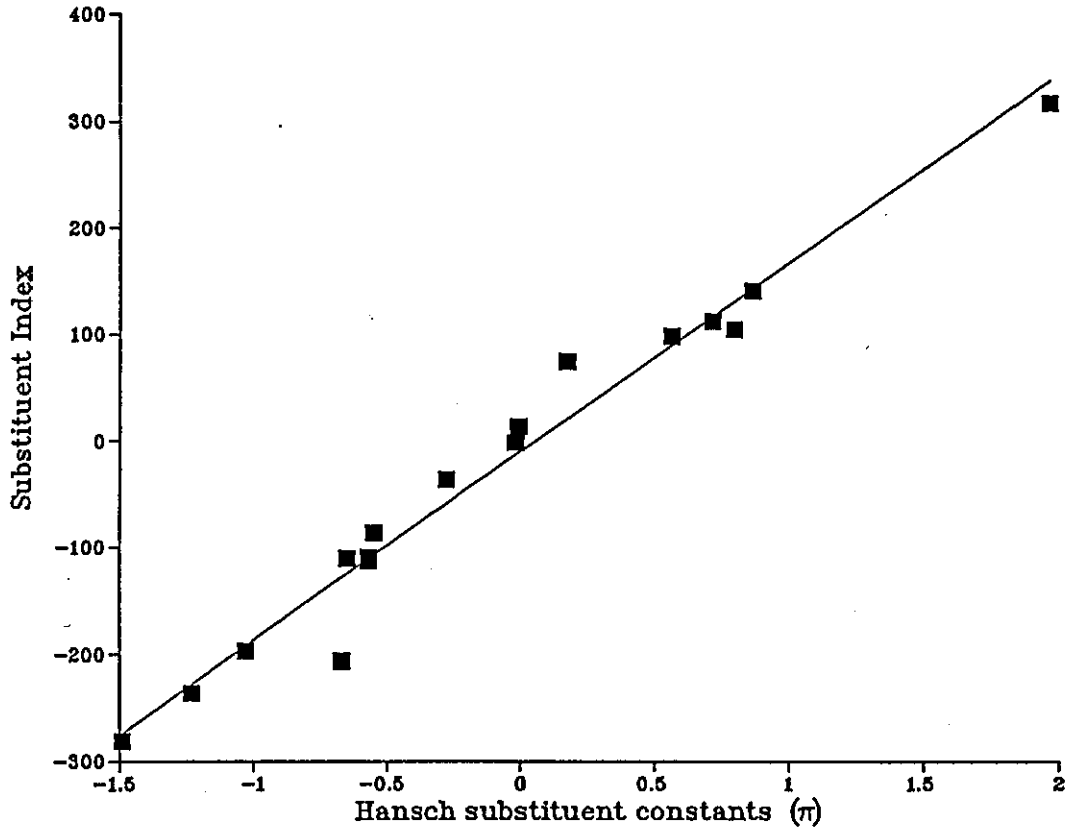
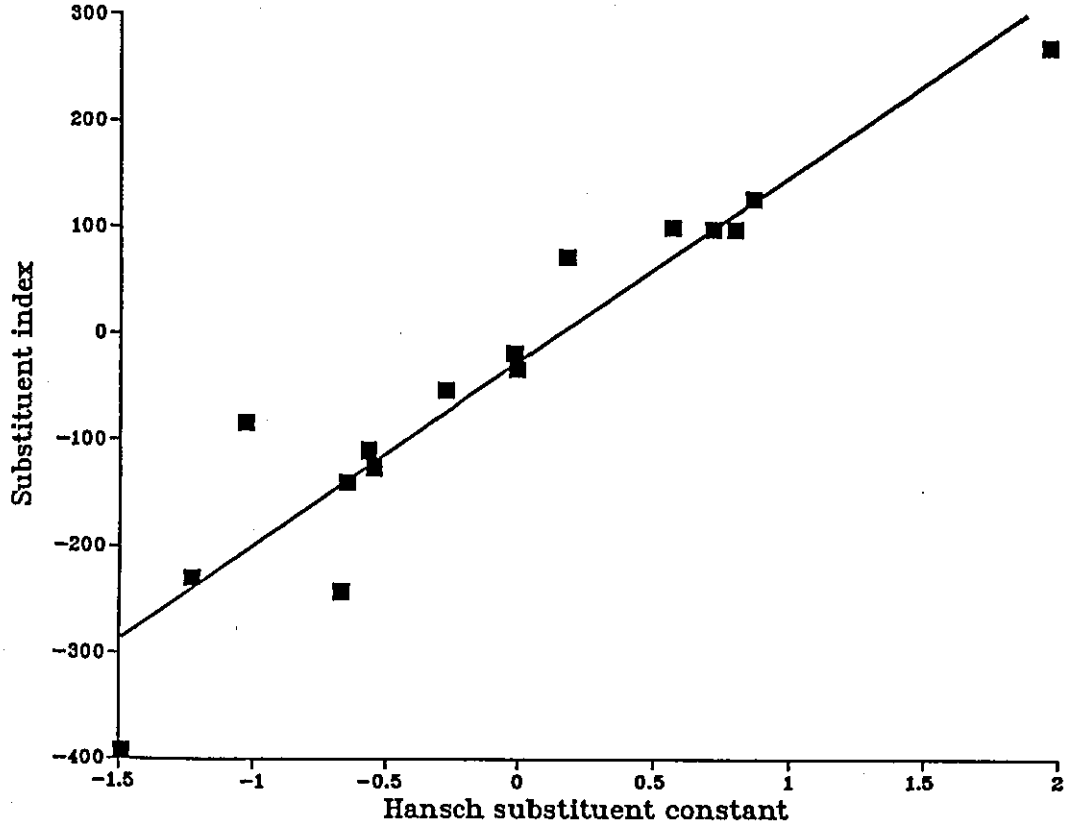


Figure 6.6: Relationship between π and substituent indices in acetonitrile eluents



CHAPTER 7

RETENTION BEHAVIOUR OF DISUBSTITUTED BENZENES

In the previous chapters (Chapter 4 and 6) the retention behaviour of mono-substituted alkylbenzenes and benzenes was examined to obtain substituent index equations. In this chapter the retention behaviour of disubstituted aromatic compounds will be examined as interactions between substituents may occur which will influence the polarity and therefore the retention of the compounds. The experimental retention indices may therefore differ from those calculated as the sum of the individual substituent contributions. In the proposed model interaction indices based on these differences will be used to account for these effects.

In most cases the substituents would not have an isolated effect on the retention index, the observed retention will be the result of the combined effects of the two substituents on the overall polarity of the molecule. The effect of each substituent will be influenced the other substituents on the ring therefore interactions terms may not be simply additive. There are various interactions which can occur between substituents on an aromatic ring, including hydrogen bonding and electronic effects involving electron donating and accepting groups. The size of these interactions will depend on both the position of substituents and their type. When substituents are in *ortho* positions steric interactions can occur and if one of the substituents is a hydroxyl group (or an amino group) there is the possibility of intra-molecular hydrogen bonding. These "through space" interactions in *ortho* substituted

compounds may be superimposed on electronic interactions but it may not be possible to separate the two effects and therefore an overall correction term will have to be derived.

The experimental interactions have been measured using two sets of model compounds. The first set of 35 compounds had a methyl group as the "fixed" substituent (i.e. $H_3CC_6H_4X$) and in the second set of 38 compounds the fixed substituent was a phenolic hydroxyl group (HOC_6H_4X). In the former set of compounds only small, relatively insignificant electronic interactions between substituents would be expected, however, steric effects could be significant with *ortho* substituted compounds. In the latter group of compounds electronic effects between substituents may have a more significant effect on the retention indices and also for appropriate functional groups intramolecular hydrogen bonding would be expected to have a major effect.

The second, variable substituents (X), were chosen from those for which individual substituent indices had previously been calculated (Chapter 6). The substituents include examples of both electron withdrawing and electron donating groups (Table 7.1). The substituents were examined in the *ortho, meta* and *para* positions relative to the fixed substituent, unless one isomer was not available (i.e. methylanisole *ortho* and *meta* only, and hydroxybenzamide *ortho* and *para* only).

The 73 compounds examined were only a small fraction of the number which would be required to study all the possible interactions between substituents. They have been used to give some idea of the size and extent of the interactions in two cases when small effects and large effects would be

Table 7.1: Classification of substituents used to obtain interaction indices according to their electron accepting and electron donating ability

From reference 177

Electron donating substituents

OH	OCH_3	CH_3
NH_2	Ph	

Electron accepting substituents

NO_2	CN	CHO
COR	Br	Cl
CO_2R	$CONH_2$	

expected.

7.1 RETENTION INDICES OF DISUBSTITUTED BENZENES

The capacity factors of the disubstituted model compounds (Tables 7.2 to 7.5) across the eluent ranges have been used to calculate the retention indices (Tables 7.6 to 7.9). For several compounds the capacity factors were less than 0.15 (marked with brackets in the tables) and the retention indices of these compounds would be expected to have a large uncertainty due to the closeness of the peak to the column void volume (Chapter 3, Section 3.2.5). Although the retention indices of these compound were calculated the values have been omitted from the regression equations used to calculate interaction indices and the analysis of results. As in the earlier chapter the retention indices with 90% organic modifier were recorded for some compounds but have been excluded form the discussion.

When the "fixed" substituent was a methyl group there were often only small differences between the retention indices of the isomers, e.g. for the bromotoluenes the difference between the isomers ranged from 3 to 15 units. In most cases the retention indices of the *meta* and *para* isomers were very similar with typical differences being less than 10 units. A difference in retention index between the *meta* and *para* isomers, and the *ortho* isomer was observed for several substituents (i.e. ester, nitro, nitrile, amide, methyl, methoxy, and phenyl) probably due to steric effects. The largest differences in retention index between isomers, with a fixed methyl group, were observed with the toluamides with a difference in retention index of up to 60 units.

Table 7.2: Capacity factors of disubstituted compounds with methyl as the fixed substituent in methanol eluents

Compound	Capacity factor (k') Methanol proportion (%)					
	40	50	60	70	80	90
2-bromotoluene	105.99	39.60	15.27	6.54	2.63	1.29
3-bromotoluene	106.98	39.79	15.39	6.48	2.59	1.25
4-bromotoluene	104.87	38.94	15.02	6.34	2.54	1.22
2-chlorotoluene	83.81	32.24	12.82	5.48	2.31	1.16
3-chlorotoluene		31.77	12.72	5.42	2.27	1.20
4-chlorotoluene	76.31	32.55	12.47	5.26	2.20	1.09
1,2-dimethylbenzene	60.72	26.61	11.34	5.11	2.25	1.15
1,3-dimethylbenzene	70.30	29.85	12.57	5.53	2.38	1.19
1,4-dimethylbenzene	72.19	30.35	13.17	5.63	2.41	1.20
2-methylacetophenone	13.35	5.76	2.92	1.55	0.87	
3-methylacetophenone	15.04	5.81	2.98	1.56	0.84	
4-methylacetophenone	14.81	5.82	2.94	1.54	0.83	
2-methylanisole	34.63	14.72	7.06	3.43	1.68	
3-methylanisole	28.90	11.76	5.80	2.92	1.47	
methyl 2-methylbenzoate	28.70	11.62	5.02	2.47	1.22	0.70
methyl 3-methylbenzoate	32.83	12.72	5.37	2.58	1.24	0.71
methyl 4-methylbenzoate	32.90	12.58	5.35	2.58	1.24	0.72
2-nitrotoluene	16.72	7.60	3.53	1.83	0.94	0.56
3-nitrotoluene	19.99	8.93	4.24	2.14	1.08	0.62
4-nitrotoluene	18.53	8.45	3.96	2.02	1.05	0.61
2-phenyltoluene	-	105.96	34.17	14.71	4.21	
3-phenyltoluene	-	125.14	38.49	13.72	4.87	
4-phenyltoluene	-	141.24	42.17	11.65	5.37	
2-tolualdehyde	10.70	4.87	2.47	1.41	0.84	
3-tolualdehyde	11.00	4.88	2.47	1.37	0.83	
4-tolualdehyde	10.16	4.57	2.33	1.32	0.81	
2-toluamide	1.43	0.97	0.49	0.39	0.28	0.23
3-toluamide	2.40	1.19	0.67	0.48	0.33	0.25
4-toluamide	2.44	1.14	0.66	0.47	0.33	0.25
2-toluidine	3.39	1.91	1.09	0.72	0.45	0.35
3-toluidine	3.52	1.98	1.10	0.71	0.44	0.34
4-toluidine	3.77	2.04	1.12	0.70	0.45	0.34
2-tolunitrile	10.81	5.28	2.32	1.29	0.71	0.48
3-tolunitrile	11.88	4.82	2.52	1.36	0.73	0.47
4-tolunitrile	11.39	5.04	2.43	1.31	0.70	0.46

Table 7.3: Capacity factors of disubstituted compounds with methyl as the fixed substituent in acetonitrile eluents

Compound	Capacity factor (k') Acetonitrile proportion (%)						
	30	40	50	60	70	80	90
2-bromotoluene	72.19	24.95	10.61	5.24	2.88	1.67	0.90
3-bromotoluene	75.32	25.59	10.59	5.20	2.83	1.62	0.86
4-bromotoluene	74.98	25.27	10.53	5.16	2.80	1.61	0.85
2-chlorotoluene	64.45	21.19	9.16	4.73	2.56	1.50	0.81
3-chlorotoluene	65.48	21.29	9.13	4.68	2.51	1.45	0.79
4-chlorotoluene	65.56	21.45	9.07	4.64	2.48	1.44	0.77
1,2-dimethylbenzene	54.74	18.27	8.16	4.16	2.34	1.37	0.75
1,3-dimethylbenzene	63.40	20.43	8.99	4.54	2.52	1.47	0.79
1,4-dimethylbenzene	63.64	20.73	9.11	4.60	2.56	1.47	0.79
2-methylacetophenone	12.59	6.02	2.96	1.77	1.07	0.76	
3-methylacetophenone	12.5	5.91	2.89	1.73	1.05	0.75	
4-methylacetophenone	11.70	5.56	3.04	1.66	1.02	0.74	
2-methylanisole	33.85	14.67	7.01	3.32	1.82	1.18	
3-methylanisole	26.81	11.94	5.84	2.84	1.58	1.04	
methyl 2-methylbenzoate	25.82	8.29	4.10	2.29	1.37	0.88	0.52
methyl 3-methylbenzoate	27.57	8.60	4.19	2.35	1.40	0.89	0.53
methyl 4-methylbenzoate	26.78	8.44	4.16	2.32	1.40	0.90	0.53
2-nitrotoluene	21.20	7.15	3.49	1.94	1.14	0.71	0.41
3-nitrotoluene	24.82	8.41	3.92	2.14	1.25	0.77	0.43
4-nitrotoluene	21.72	7.67	3.74	2.06	1.21	0.75	0.42
2-phenyltoluene		77.11	22.18	9.09	4.12	2.31	
3-phenyltoluene		83.16	23.70	9.60	4.32	2.40	
4-phenyltoluene		87.94	24.83	10.02	4.49	2.49	
2-tolualdehyde	10.73	5.43	2.89	1.70	1.11	0.58	
3-tolualdehyde	11.26	5.62	2.98	1.73	1.14	0.58	
4-tolualdehyde	10.18	5.15	2.77	1.65	1.10	0.57	
2-toluamide	1.35	0.74	0.52	0.41	0.33	0.34	0.25
3-toluamide	1.70	0.92	0.62	0.46	0.36	0.33	0.27
4-toluamide	1.90	1.01	0.61	0.47	0.36	0.35	0.28
2-toluidine	4.65	2.30	1.44	0.97	0.65	0.48	0.32
3-toluidine	4.87	2.33	1.43	0.96	0.64	0.47	0.29
4-toluidine	4.85	2.30	1.43	0.96	0.66	0.48	0.33
2-tolunitrile	11.43	5.08	2.67	1.58	0.98	0.65	0.40
3-tolunitrile	13.03	5.65	2.91	1.68	1.04	0.66	0.40
4-tolunitrile	12.21	5.33	2.77	1.61	0.99	0.65	0.39

Table 7.4: Capacity factors of disubstituted compounds with hydroxyl as the fixed substituent in methanol eluents

Compound	Capacity factor (k') Methanol proportion (%)					
	40	50	60	70	80	90
2-aminophenol	0.78	0.66	0.42	0.30	0.26	0.25
3-aminophenol	0.35	0.28	0.25	0.19	0.18	0.20
4-aminophenol	0.25	0.22	0.19	0.16	0.21	0.21
2-bromophenol	6.12	3.34	1.70	0.86	0.45	0.26
3-bromophenol	9.38	4.82	2.17	1.11	0.56	0.32
4-bromophenol	8.80	4.56	2.14	1.10	0.57	0.34
2-chlorophenol	4.86	2.72	1.36	0.74	0.42	0.25
3-chlorophenol	7.38	3.95	1.85	0.93	0.51	0.30
4-chlorophenol	6.94	3.65	1.76	0.91	0.51	0.32
1,2-dihydroxybenzene	0.96	0.64	0.43	0.31	0.25	0.26
1,3-dihydroxybenzene	0.53	0.37	0.27	0.19	0.18	0.21
1,4-dihydroxybenzene	0.31	0.27	0.20	0.18	0.18	0.22
2-hydroxyacetophenone	9.02	4.33	2.21	1.25	0.84	0.55
3-hydroxyacetophenone	2.02	1.02	0.58	0.39	0.34	0.30
4-hydroxyacetophenone	1.55	0.78	0.44	0.29	0.19	(0.11)
2-hydroxybenzaldehyde	4.99	2.52	1.41	0.83	0.53	
3-hydroxybenzaldehyde	1.64	0.89	0.54	0.37	0.25	
4-hydroxybenzaldehyde	1.00	0.51	0.27	(0.14)	(0.07)	
2-hydroxybenzamide	1.57	0.84	0.50	0.31	0.20	
4-hydroxybenzamide	0.32	0.22	0.16	(0.13)	(0.11)	
2-hydroxybenzonitrile	1.31	0.74	0.34	0.18	-	-
3-hydroxybenzonitrile	2.17	1.26	0.68	0.42	0.21	(0.10)
4-hydroxybenzonitrile	1.48	0.89	0.47	0.26	(0.07)	-
2-methoxyphenol	2.95	1.49	0.92	0.59	0.37	
3-methoxyphenol	2.63	1.29	0.78	0.50	0.31	
4-methoxyphenol	1.93	0.99	0.63	0.42	0.29	
methyl 2-hydroxybenzoate	19.60	10.09	4.65	4.66	1.16	0.70
methyl 3-hydroxybenzoate	3.72	1.96	0.99	0.58	0.34	0.25
methyl 4-hydroxybenzoate	3.39	1.88	0.88	0.49	0.25	(0.13)
2-methylphenol	4.75	2.76	1.42	0.78	0.47	0.33
3-methylphenol	4.41	2.44	1.29	0.72	0.44	0.32
4-methylphenol	4.86	2.51	1.31	0.73	0.45	0.32
2-nitrophenol	3.56	1.96	0.91	0.40	(0.08)	
3-nitrophenol	3.38	1.96	1.00	0.53	0.23	(0.09)
4-nitrophenol	1.27	0.69	0.29	(0.13)	-	
2-phenylphenol	33.87	10.73	4.48	1.97	0.90	
3-phenylphenol	35.87	11.35	4.53	1.92	0.92	
4-phenylphenol	35.74	11.52	4.12	1.98	0.96	

Table 7.5: Capacity factors of disubstituted compounds with hydroxyl as the fixed substituent in acetonitrile eluents

Compound	Capacity factor (k') Acetonitrile proportion (%)						
	30	40	50	60	70	80	90
2-aminophenol	1.08	0.79	0.63	0.54	0.42	0.44	0.35
3-aminophenol	0.63	0.52	0.38	0.30	0.23	0.26	0.18
4-aminophenol	0.40	0.38	0.29	0.24	0.20	0.26	0.22
2-bromophenol	5.82	3.06	1.6	0.97	0.61	0.45	0.27
3-bromophenol	7.85	3.77	1.84	1.08	0.66	0.48	0.30
4-bromophenol	7.49	3.60	1.77	1.05	0.65	0.49	0.30
2-chlorophenol	4.69	2.75	1.43	0.89	0.56	0.44	0.26
3-chlorophenol	6.98	3.26	1.64	0.99	0.61	0.47	0.28
4-chlorophenol	6.39	3.04	1.57	0.94	0.59	0.45	0.29
1,2-dihydroxybenzene	0.85	0.78	0.53	0.42	0.31	0.41	0.25
1,3-dihydroxybenzene	0.55	0.55	0.38	0.29	0.22	0.25	(0.14)
1,4-dihydroxybenzene	0.35	0.43	0.32	0.25	0.21	0.24	(0.14)
2-hydroxyacetophenone	7.57	4.51	2.16	1.52	0.84	0.65	0.42
3-hydroxyacetophenone	1.70	1.31	0.67	0.55	0.29	0.32	0.20
4-hydroxyacetophenone	1.15	0.88	0.53	0.46	0.27	0.29	0.23
2-hydroxybenzaldehyde	6.86	5.56	2.39	1.34	0.94	0.56	
3-hydroxybenzaldehyde	1.77	1.11	0.74	0.50	0.35	0.19	
4-hydroxybenzaldehyde	1.20	0.78	0.54	0.40	0.29	0.16	
2-hydroxybenzamide	1.52	0.95	0.76	0.47	0.35	0.21	
4-hydroxybenzamide	0.27	0.23	0.19	0.17	(0.14)	(0.01)	
2-hydroxybenzonitrile	1.51	1.02	0.66	0.48	0.30	0.22	0.19
3-hydroxybenzonitrile	2.50	1.47	0.87	0.58	0.38	0.33	0.23
4-hydroxybenzonitrile	1.74	1.18	0.73	0.49	0.34	0.26	(0.15)
2-methoxyphenol	3.12	1.90	1.12	0.77	0.51	0.41	
3-methoxyphenol	2.73	1.62	0.97	0.64	0.42	0.32	
4-methoxyphenol	2.05	1.30	0.82	0.56	0.38	0.29	
methyl 2-hydroxybenzoate	14.59	7.29	3.41	2.00	1.18	0.82	0.45
methyl 3-hydroxybenzoate	3.10	1.67	0.95	0.61	0.41	0.36	0.23
methyl 4-hydroxybenzoate	2.44	1.48	0.88	0.61	0.40	0.33	0.18
2-methylphenol	5.26	2.55	1.41	0.89	0.58	0.41	0.26
3-methylphenol	4.45	2.28	1.28	0.81	0.53	0.25	0.25
4-methylphenol	5.10	2.28	1.28	0.82	0.54	0.24	0.27
2-nitrophenol	5.27	2.59	1.57	1.07	0.65	0.36	0.29
3-nitrophenol	4.23	2.03	1.12	0.71	0.45	0.35	0.24
4-nitrophenol	2.12	1.06	0.67	0.49	0.28	0.19	0.17
2-phenylphenol	31.39	10.51	4.15	1.98	1.04	0.67	
3-phenylphenol	27.91	9.01	3.63	1.76	0.93	0.62	
4-phenylphenol	27.59	8.86	3.36	1.72	0.99	0.63	

Table 7.6: Retention indices of disubstituted compounds with methyl as the fixed substituent in methanol eluents

Compound	Retention index Methanol proportion (%)					
	40	50	60	70	80	90
2-bromotoluene	1134	1157	1179	1205	1244	1296
3-bromotoluene	1135	1157	1180	1203	1240	1282
4-bromotoluene	1132	1154	1176	1199	1235	1270
2-chlorotoluene	1107	1129	1150	1174	1207	1248
3-chlorotoluene		1127	1149	1172	1202	1232
4-chlorotoluene	1096	1130	1144	1166	1192	1223
1,2-dimethylbenzene	1071	1103	1130	1159	1199	1245
1,3-dimethylbenzene	1088	1118	1147	1179	1215	1262
1,2-dimethylbenzene	1091	1121	1155	1180	1219	1264
2-methylacetophenone	894	896	895	898	902	
3-methylacetophenone	900	900	899	900	903	
4-methylacetophenone	898	900	897	896	899	
2-methylanisole	1003	1026	1044	1063	1090	
3-methylanisole	974	996	1011	1030	1055	
methyl 2-methylbenzoate	984	991	997	999	1019	1025
methyl 3-methylbenzoate	1000	1003	1008	1008	1025	1029
methyl 4-methylbenzoate	1000	1001	1007	1009	1025	1033
2-nitrotoluene	922	933	938	942	945	923
3-nitrotoluene	942	955	969	976	985	970
4-nitrotoluene	934	948	957	963	977	965
2-phenyltoluene		1296	1303	1318	1340	
3-phenyltoluene		1322	1336	1352	1379	
4-phenyltoluene		1334	1351	1367	1397	
2-tolualdehyde	869	873	877	881	894	
3-tolualdehyde	872	873	876	876	890	
4-tolualdehyde	863	865	866	868	883	
2-toluamide	628	655	617	609	589	524
3-toluamide	688	682	668	654	638	575
4-toluamide	689	677	664	651	639	567
2-toluidine	728	747	747	742	730	710
3-toluidine	732	751	748	739	731	701
4-toluidine	740	755	750	745	722	700
2-tolunitrile	871	872	870	863	860	856
3-tolunitrile	882	884	884	873	869	849
4-tolunitrile	877	878	878	865	857	831

Table 7.7: Retention indices of disubstituted compounds with methyl as the fixed substituent in acetonitrile eluents

Compound	Retention index Acetonitrile proportion (%)						
	30	40	50	60	70	80	90
2-bromotoluene	1140	1152	1168	1183	1202	1223	1260
3-bromotoluene	1145	1156	1167	1181	1197	1211	1242
4-bromotoluene	1145	1154	1166	1178	1194	1208	1236
2-chlorotoluene	1115	1126	1138	1151	1168	1182	1215
3-chlorotoluene	1117	1126	1137	1149	1162	1171	1201
4-chlorotoluene	1117	1128	1136	1146	1159	1168	1192
1,2-dimethylbenzene	1085	1102	1115	1125	1141	1150	1179
1,3-dimethylbenzene	1104	1120	1134	1147	1163	1174	1201
1,4-dimethylbenzene	1104	1122	1137	1150	1168	1176	1204
2-methylacetophenone	888	887	886	889	892	892	
3-methylacetophenone	887	884	881	883	886	887	
4-methylacetophenone	879	875	873	874	879	882	
2-methylanisole	1014	1024	1030	1037	1043	1046	
3-methylanisole	985	992	996	1001	1003	1002	
methyl 2-methylbenzoate	976	974	976	977	985	985	1009
methyl 3-methylbenzoate	984	980	980	984	991	991	1018
methyl 4-methylbenzoate	981	977	979	980	990	993	1020
2-nitrotoluene	952	950	943	935	932	905	898
3-nitrotoluene	971	976	967	960	959	937	929
4-nitrotoluene	955	961	957	950	949	926	919
2-phenyltoluene		1274	1273	1274	1276	1277	
3-phenyltoluene		1286	1286	1287	1289	1291	
4-phenyltoluene		1294	1295	1297	1300	1304	
2-tolualdehyde	868	871	870	880	893	879	
3-tolualdehyde	873	877	876	884	901	877	
4-tolualdehyde	861	863	862	873	890	872	
2-toluamide	621	584	561	550	568	631	683
3-toluamide	650	619	594	581	597	625	706
4-toluamide	664	633	592	582	594	642	724
2-toluidine	765	766	764	762	769	761	785
3-toluidine	771	769	763	761	762	754	746
4-toluidine	770	767	763	761	770	760	799
2-tolunitrile	898	895	889	884	886	874	887
3-tolunitrile	914	912	906	900	905	880	890
4-tolunitrile	906	903	897	890	891	871	874

Table 7.8: Retention indices of disubstituted compounds with hydroxyl as the fixed substituent in methanol eluents

Compound	Retention index Methanol proportion (%)					
	40	50	60	70	80	90
2-aminophenol	576	592	580	562	569	561
3-aminophenol	483	477	492	467	474	476
4-aminophenol	446	446	449	429	515	487
2-bromophenol	813	810	809	777	734	592
3-bromophenol	862	860	849	830	793	680
4-bromophenol	855	852	847	829	799	703
2-chlorophenol	786	783	772	753	708	562
3-chlorophenol	834	833	823	803	767	643
4-chlorophenol	827	822	815	798	768	683
1,2-dihydroxybenzene	599	589	583	570	568	582
1,3-dihydroxybenzene	531	516	506	483	476	493
1,4-dihydroxybenzene	469	474	461	451	467	502
2-hydroxyacetophenone	853	860	867	875	888	917
3-hydroxyacetophenone	679	661	641	619	614	612
4-hydroxyacetophenone	649	624	594	551	428	(118)
2-hydroxybenzaldehyde	782	783	783	769	771	
3-hydroxybenzaldehyde	656	641	623	600	571	
4-hydroxybenzaldehyde	600	565	507	(396)	(243)	
2-hydroxybenzamide	651	636	605	563	518	
4-hydroxybenzamide	463	454	418	(384)	(350)	
2-hydroxybenzonitrile	634	608	547	447		
3-hydroxybenzonitrile	693	678	658	625	510	(186)
4-hydroxybenzonitrile	649	632	599	529	(196)	
2-methoxyphenol	715	711	705	698	682	
3-methoxyphenol	701	692	678	663	635	
4-methoxyphenol	666	656	643	628	606	
methyl 2-hydroxybenzoate	947	959	975	989	1006	1027
methyl 3-hydroxybenzoate	755	738	720	700	654	565
methyl 4-hydroxybenzoate	744	733	701	666	565	(281)
2-methylphenol	783	785	779	766	746	695
3-methylphenol	775	768	764	748	725	672
4-methylphenol	777	772	766	752	733	682
2-nitrophenol	750	739	706	622	(228)	
3-nitrophenol	744	739	721	685	543	(83)
4-nitrophenol	631	597	519	(379)		
2-phenylphenol	992	981	969	947	922	
3-phenylphenol	1007	989	973	943	918	
4-phenylphenol	1006	991	970	949	929	

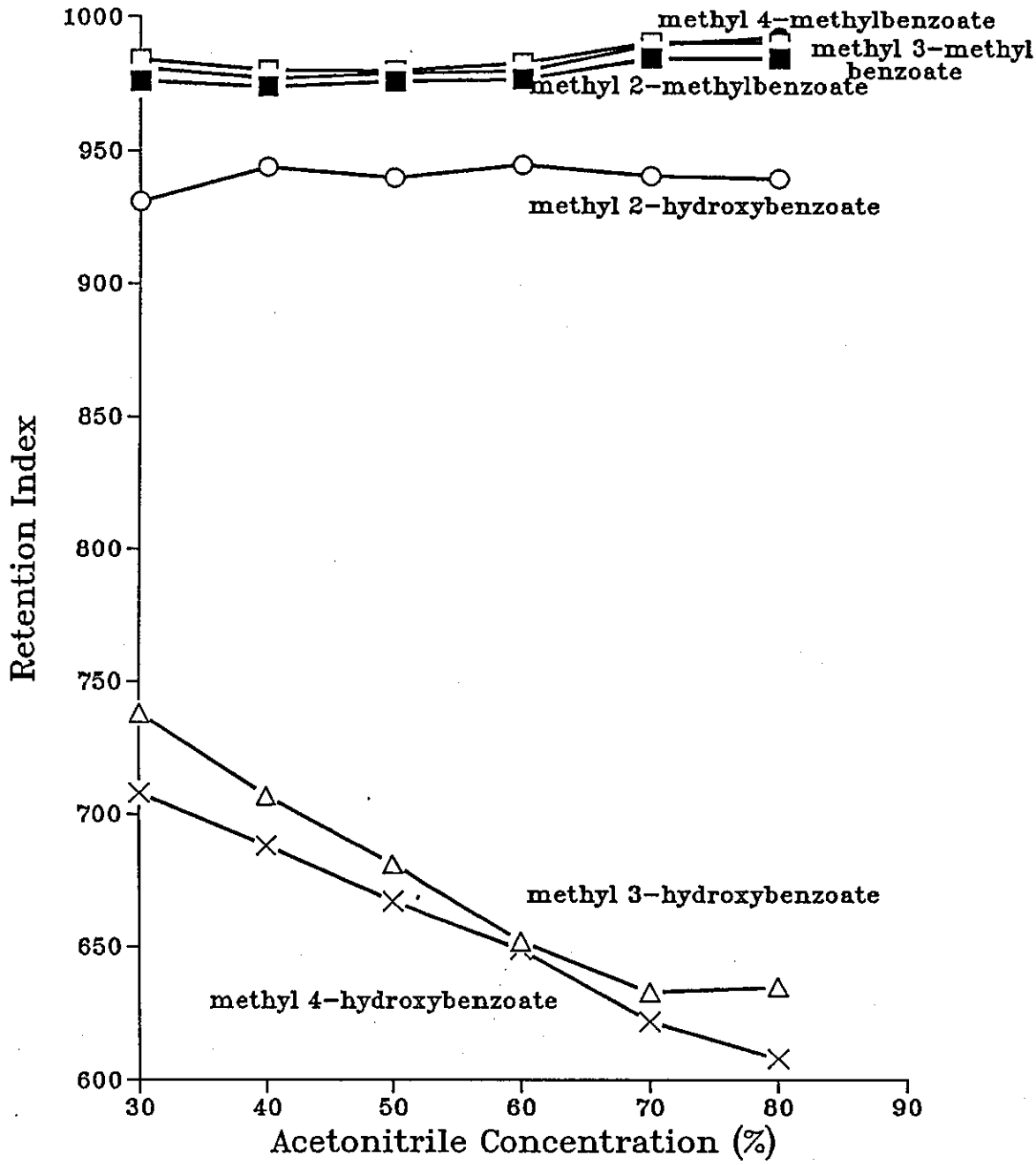
Table 7.9: Retention indices of disubstituted compounds with hydroxyl as the fixed substituent in acetonitrile eluents

Compound	Retention index Acetonitrile proportion (%)						
	30	40	50	60	70	80	90
2-aminophenol	586	589	600	618	635	711	830
3-aminophenol	519	522	496	474	468	519	520
4-aminophenol	463	471	445	423	424	516	610
2-bromophenol	821	805	786	765	749	723	712
3-bromophenol	859	838	815	792	772	743	752
4-bromophenol	853	831	807	785	765	747	749
2-chlorophenol	786	788	764	744	724	714	681
3-chlorophenol	836	815	792	769	749	731	724
4-chlorophenol	825	804	783	758	739	717	733
1,2-dihydroxybenzene	577	587	565	558	553	687	671
1,3-dihydroxybenzene	522	532	498	467	453	495	(409)
1,4-dihydroxybenzene	467	492	461	433	435	493	(394)
2-hydroxyacetophenone	849	867	853	854	855	857	878
3-hydroxyacetophenone	663	669	616	605	588	597	499
4-hydroxyacetophenone	615	606	569	562	556	562	462
2-hydroxybenzaldehyde	862	875	834	822	845	867	
3-hydroxybenzaldehyde	648	630	610	592	571	510	
4-hydroxybenzaldehyde	600	577	550	537	513	462	
2-hydroxybenzamide	629	608	616	577	568	540	
4-hydroxybenzamide	420	393	356	336	(324)	(251)	
2-hydroxybenzonitrile	648	630	608	589	542	463	554
3-hydroxybenzonitrile	711	688	665	639	610	606	630
4-hydroxybenzonitrile	463	471	445	423	424	516	610
2-methoxyphenol	720	712	697	693	682	678	
3-methoxyphenol	704	688	668	653	625	593	
4-methoxyphenol	668	654	636	622	601	565	
methyl 2-hydroxybenzoate	931	944	940	945	941	940	943
methyl 3-hydroxybenzoate	738	707	681	652	633	635	629
methyl 4-hydroxybenzoate	708	688	667	649	622	608	524
2-methylphenol	788	776	761	745	732	714	691
3-methylphenol	772	758	741	722	708	691	665
4-methylphenol	774	758	741	723	713	700	707
2-nitrophenol	781	778	783	789	766	641	739
3-nitrophenol	754	739	716	688	659	623	409
4-nitrophenol	669	635	611	598	518	395	394
2-phenylphenol	984	973	939	916	884	847	
3-phenylphenol	985	950	914	888	854	822	
4-phenylphenol	984	947	910	879	853	816	

When the "fixed" substituent was the phenolic hydroxyl there were many more differences between the retention indices of isomers and the effects were often very large. For a few combinations the *meta* and *para* isomers were similar (phenyl, chloro, bromo, and methyl) but in these the cases the *ortho* isomer usually differed by a larger amount than in substituted toluene compounds (e.g. the bromotoluenes had no significant difference between isomers but the bromophenol, *ortho* isomer was 40 - 65 less than the *meta* and *para* isomers). The largest differences between isomers were observed when intra-molecular hydrogen bonding could occur in the *ortho* substituted compounds, such as phenols substituted with carbonyl containing groups (-CH=O, -C(R)=O, -C(OR)=O, and C(NH_2)=O) . In these compounds the retention index of the *ortho* isomer was considerably larger than that of the other isomers (e.g. the retention index of methyl 2-hydroxybenzoate was about 200 - 300 units higher than the 3- and 4- isomers, Figure 7.1 which contrasts the methyl methylbenzoate and methyl hydroxybenzoate isomers).

In a few cases there appear to be specific cross-ring electronic effects in which the *para* isomer differed considerably from the *meta* isomer. This was particularly apparent with the hydroxybenzonitriles, where the *para* isomer was eluted considerably earlier than the *meta* isomer (90 - 250 units in acetonitrile and 50 - 90 units in methanol). A similar effect may be present in the nitrophenols, however, as the pKa of 4-nitrophenol is 7.15[165], this compound may be partially ionised at the working pH which would result in large differences between the isomers.

Figure 7.1: Retention indices of isomers of methyl methylbenzoate and methyl hydroxybenzoate in acetonitrile eluents



7.2 INTERACTION INCREMENTS FOR PAIRS OF AROMATIC SUBSTITUENTS

To calculate the size of any interactions between substituents the summation of the parent index and substituent indices ($RI_{SUM} = PI + \Sigma SI + SI_R$) was calculated (Table 7.10 and 7.11). The difference between the experimental retention indices and the summation has been used to calculate the interaction increments (δI) using the equation

$$\delta I = RI_{exp} - (PI + \Sigma SI + SI_R) \qquad 7.1$$

at each eluent composition (Tables 7.12 to 7.15). The substituent index equations and parent index equations cannot reliably be used for extrapolation (see Chapter 4 and 6), therefore the calculated values were limited to the eluent ranges 40 - 80% methanol and 30 - 80% acetonitrile. Like the substituent retention index increments these increments can be converted into interaction index equations and the coefficients for the effect of modifier determined (see later). However, as each pair of substituents will generate a separate equation the very large number of potential combinations will present problems for a database. Any attempt to use an expert system to access the data will be complicated as each substituent combination will also require a specific rule for its recognition.

The simple approach of an equation for each pair of substituents will not therefore be practical in the long term and it will be important in future work to try and discover the underlying rules which govern the effects so

Table 7.10: Retention indices calculated as the sum of parent index and substituent indices

SUM	RI_{SUM} Methanol proportion (%) 40	50	60	70	80
$PI+SI_{CH3}+SI_{Br}$	1128	1147	1167	1188	1209
$PI+SI_{CH3}+SI_{Cl}$	1098	1117	1136	1153	1170
$PI+SI_{CH3}+SI_{CH3}$	1084	1112	1138	1161	1182
$PI+SI_{CH3}+SI_{COR}+SI_{CH3}$	899	899	899	899	899
$PI+SI_{CH3}+SI_{OR}+SI_{CH3}$	984	1001	1018	1036	1053
$PI+SI_{CH3}+SI_{CO2R}+SI_{CH3}$	999	1002	1006	1010	1014
$PI+SI_{CH3}+SI_{NO2}$	950	958	965	971	975
$PI+SI_{CH3}+SI_{Ph}$	1305	1317	1332	1348	1368
$PI+SI_{CH3}+SI_{CHO}$	874	874	876	878	882
$PI+SI_{CH3}+SI_{CONH2}$	703	691	679	666	653
$PI+SI_{CH3}+SI_{NH2}$	748	758	760	755	741
$PI+SI_{CH3}+SI_{CN}$	875	878	877	871	860
$PI+SI_{OH}+SI_{NH2}$	443	429	403	363	310
$PI+SI_{OH}+SI_{Br}$	822	818	810	796	778
$PI+SI_{OH}+SI_{Cl}$	792	788	778	762	739
$PI+SI_{OH}+SI_{OH}$	473	454	423	378	319
$PI+SI_{OH}+SI_{CHO}$	568	545	518	487	451
$PI+SI_{OH}+SI_{COR}+SI_{CH3}$	593	570	542	508	468
$PI+SI_{OH}+SI_{CONH2}$	397	362	321	275	222
$PI+SI_{OH}+SI_{CN}$	569	549	519	479	429
$PI+SI_{OH}+SI_{OR}+SI_{CH3}$	679	672	661	644	622
$PI+SI_{OH}+SI_{CO2R}+SI_{CH3}$	693	673	648	618	583
$PI+SI_{OH}+SI_{CH3}$	778	783	780	769	750
$PI+SI_{OH}+SI_{NO2}$	644	629	608	579	544
$PI+SI_{OH}+SI_{Ph}$	1000	988	974	957	937

Table 7.11: Retention Indices calculated as sum of parent index and substituent indices

SUM	RI_{SUM} Acetonitrile proportion (%)					
	30	40	50	60	70	80
$PI+SI_{CH3}+SI_{Br}$	1137	1154	1168	1178	1186	1190
$PI+SI_{CH3}+SI_{Cl}$	1108	1125	1139	1149	1157	1161
$PI+SI_{CH3}+SI_{CH3}$	1110	1127	1141	1151	1159	1163
$PI+SI_{CH3}+SI_{COR}+SI_{CH3}$	900	900	900	900	900	900
$PI+SI_{CH3}+SI_{OR}+SI_{CH3}$	1000	1008	1013	1016	1016	1014
$PI+SI_{CH3}+SI_{CO2R}+SI_{CH3}$	990	993	995	996	997	996
$PI+SI_{CH3}+SI_{NO2}$	970	973	971	964	953	936
$PI+SI_{CH3}+SI_{Ph}$	1300	1298	1296	1295	1294	1295
$PI+SI_{CH3}+SI_{CHO}$	880	885	888	888	886	881
$PI+SI_{CH3}+SI_{CONH2}$	682	637	614	614	636	679
$PI+SI_{CH3}+SI_{CN}$	913	916	915	910	902	890
$PI+SI_{OH}+SI_{NH2}$	484	454	428	405	386	371
$PI+SI_{OH}+SI_{Br}$	825	811	798	787	777	768
$PI+SI_{OH}+SI_{Cl}$	796	782	769	758	748	739
$PI+SI_{OH}+SI_{OH}$	486	441	402	368	340	318
$PI+SI_{OH}+SI_{CHO}$	568	543	519	497	476	458
$PI+SI_{OH}+SI_{COR}+SI_{CH3}$	588	557	531	509	492	479
$PI+SI_{OH}+SI_{CONH2}$	367	293	245	223	228	259
$PI+SI_{OH}+SI_{CN}$	601	573	546	519	493	468
$PI+SI_{OH}+SI_{OR}+SI_{CH3}$	688	665	644	624	607	591
$PI+SI_{OH}+SI_{CO2R}+SI_{CH3}$	678	650	626	605	587	573
$PI+SI_{OH}+SI_{CH3}$	798	784	771	760	750	741
$PI+SI_{OH}+SI_{NO2}$	657	630	602	573	543	513
$PI+SI_{OH}+SI_{Ph}$	989	955	927	903	885	872

Table 7.12: Interactions between substituents with the fixed substituent as methyl in methanol eluents

Substituent pairs	Interaction increment Methanol proportion (%)				
	40	50	60	70	80
CH_3 + 2-Br	6	10	12	17	35
CH_3 + 3-Br	7	10	13	15	31
CH_3 + 4-Br	4	7	9	11	26
CH_3 + 2-Cl	9	12	14	21	37
CH_3 + 3-Cl	4	10	13	19	32
CH_3 + 4-Cl	-2	13	8	13	22
CH_3 + 2-CH_3	-13	-9	-8	-2	17
CH_3 + 3-CH_3	4	6	9	18	33
CH_3 + 4-CH_3	7	9	17	19	37
CH_3 + 2-$COCH_3$	-6	-4	-5	-2	2
CH_3 + 3-$COCH_3$	0	0	-1	0	3
CH_3 + 4-$COCH_3$	-2	0	-3	-4	0
CH_3 + 2-OCH_3	19	25	26	27	37
CH_3 + 3-OCH_3	-10	-5	-7	-6	2
CH_3 + 2-CO_2CH_3	-15	-11	-9	-11	5
CH_3 + 3-CO_2CH_3	1	1	-2	-2	11
CH_3 + 4-CO_2CH_3	1	-1	-1	-1	11
CH_3 + 2-NO_2	-28	-25	-27	-29	-30
CH_3 + 3-NO_2	-8	-3	4	5	10
CH_3 + 4-NO_2	-16	-10	-8	-14	2
CH_3 + 2-Ph		-21	-29	-30	-28
CH_3 + 3-Ph		5	4	4	11
CH_3 + 4-Ph		17	19	19	29
CH_3 + 2-CHO	-5	-1	1	3	12
CH_3 + 3-CHO	-2	-1	0	-2	8
CH_3 + 4-CHO	-11	-9	-10	-10	1
CH_3 + 2-$CONH_2$	-75	-36	-62	-57	-64
CH_3 + 3-$CONH_2$	-15	-9	-11	-12	-15
CH_3 + 4-$CONH_2$	-14	-14	-9	-15	-14
CH_3 + 2-NH_2	-20	-11	-12	-13	-11
CH_3 + 3-NH_2	-16	-7	-12	-14	-9
CH_3 + 4-NH_2	-8	-3	-10	-10	-19
CH_3 + 2-CN	-4	-6	-7	-8	0
CH_3 + 3-CN	7	6	7	2	9
CH_3 + 4-CN	2	0	1	-6	-3

Table 7.13: Interactions between substituents with the fixed substituent as methyl in acetonitrile eluents

Substituent pairs	Interaction increment Acetonitrile proportion (%)					
	30	40	50	60	70	80
CH_3 + 2-Br	3	-2	0	5	16	33
CH_3 + 3-Br	8	2	-1	3	13	21
CH_3 + 4-Br	8	0	-2	0	8	18
CH3 + 2-Cl	7	1	-1	2	11	21
CH_3 + 3-Cl	9	1	-2	0	5	10
CH_3 + 4-Cl	9	3	-3	-3	2	7
CH_3 + 2-CH_3	-25	-25	-26	-26	-18	-13
CH_3 + 3-CH_3	-6	-7	-7	-4	4	11
CH_3 + 4-CH3	-6	-5	-4	-1	9	13
CH_3 + 2-$COCH_3$	-12	-13	-14	-11	-8	-8
CH_3 + 3-$COCH_3$	-13	-16	-19	-17	-14	-13
CH_3 + 4-$COCH_3$	-21	-25	-27	-26	-21	-18
CH_3 + 2-OCH_3	14	16	17	21	27	32
CH_3 + 3-OCH_3	-15	-16	-17	-15	-13	-12
CH_3 + 2-CO_2CH_3	-14	-19	-19	-19	-12	-11
CH_3 + 3-CO_2CH_3	-6	-13	-15	-12	-6	-5
CH_3 + 4-CO_2CH_3	-9	-16	-16	-16	-7	-3
CH_3 + 2-NO_2	-18	-23	-28	-29	-21	-31
CH_3 + 3-NO_2	1	3	-4	-4	6	1
CH_3 + 4-NO_2	-15	-12	-14	-14	-4	-10
CH_3 + 2-Ph		-24	-23	-21	-18	-18
CH_3 + 3-Ph		-12	-10	-8	-5	-4
CH_3 + 4-Ph		-4	-1	2	6	9
CH_3 + 2-CHO	-12	-14	-18	-8	7	-2
CH_3 + 3-CHO	-7	-8	-12	-4	15	-4
CH_3 + 4-CHO	-19	-22	-26	-15	4	-9
CH_3 + 2-$CONH_2$	-61	-53	-53	-64	-68	-48
CH_3 + 3-$CONH_2$	-32	-18	-20	-33	-39	-54
CH_3 + 4-$CONH_2$	-18	-4	-22	-32	-42	-37
CH_3 + 2-NH_2	-31	-31	-33	-34	-26	-32
CH_3 + 3-NH_2	-25	-32	-34	-35	-33	-49
CH_3 + 4-NH_2	-26	-30	-34	-35	-25	-33
CH_3 + 2-CN	-15	-21	-26	-26	-16	-16
CH_3 + 3-CN	1	-4	-9	-10	3	-10
CH_3 + 4-CN	-7	-13	-18	-20	-11	-19

Table 7.14: Interactions between substituents where the fixed substituent was phenolic hydroxyl in methanol eluents

Substituent pairs	Interaction increment Methanol proportion (%)				
	40	50	60	70	80
OH + 2-NH_2	133	163	177	199	259
OH + 3-NH_2	40	48	89	104	164
OH + 4-NH_2	3	17	46	66	205
OH + 2-Br	-9	-8	-1	-19	-44
OH + 3-Br	40	42	39	34	15
OH + 4-Br	33	34	37	33	21
OH + 2-Cl	-6	-5	-6	-9	-31
OH + 3-Cl	42	45	45	41	28
OH + 4-Cl	35	34	37	36	29
OH + 2-OH	126	135	160	192	249
OH + 3-OH	58	62	83	105	157
OH + 4-OH	-4	20	37	73	148
OH + 2-CHO	214	238	265	282	320
OH + 3-CHO	88	96	105	113	120
OH + 4-CHO	32	20	-11	(-91)	(-208)
OH + 2-$COCH_3$	260	290	325	367	420
OH + 3-$COCH_3$	86	91	99	111	145
OH + 4-$COCH_3$	56	55	52	43	-40
OH + 2-$CONH_2$	254	274	284	288	296
OH + 4-$CONH_2$	66	92	97	(109)	(128)
OH + 2-CN	65	59	28	-32	
OH + 3-CN	124	129	139	146	81
OH + 4-CN	80	83	80	50	(-233)
OH + 2-OCH_3	36	39	44	54	60
OH + 3-OCH_3	22	20	17	19	13
OH + 4-OCH_3	-13	-16	-18	-16	-16
OH + 2-CO_2CH_3	254	286	327	371	423
OH + 3-CO_2CH_3	62	65	72	82	71
OH + 4-CO_2CH_3	51	60	53	48	-18
OH + 2-CH_3	5	2	-1	-3	-4
OH + 3-CH_3	-3	-15	-16	-21	-25
OH + 4-CH_3	-1	-11	-14	-17	-17
OH + 2-NO_2	106	110	98	43	(-316)
OH + 3-NO_2	100	110	113	106	-1
OH + 4-NO_2	-13	-32	-89	(-200)	
OH + 2-Ph	-8	2	-1	-3	-4
OH + 3-Ph	-3	-15	-16	-21	-25
OH + 4-Ph	-1	-11	-14	-17	-17

Table 7.15: Interactions between substituents where the fixed substituent was phenolic hydroxyl in acetonitrile eluents

Substituent pairs	Interaction increment Acetonitrile proportion (%)					
	30	40	50	60	70	80
OH + 2-NH_2	102	135	172	213	249	340
OH + 3-NH_2	35	68	68	69	82	148
OH + 4-NH_2	-21	17	17	18	38	145
OH + 2-Br	-4	-6	-12	-22	-28	-45
OH + 3-Br	34	27	17	5	-5	-25
OH + 4-Br	28	20	9	-2	-12	-21
OH + 2-Cl	-10	6	-5	-14	-24	-25
OH + 3-Cl	40	33	23	11	1	-8
OH + 4-Cl	29	22	14	0	-9	-22
OH + 2-OH	91	146	163	190	213	346
OH + 3-OH	36	91	96	99	113	177
OH + 4-OH	-19	51	59	65	95	175
OH + 2-CHO	294	332	315	325	369	409
OH + 3-CHO	80	87	91	95	95	52
OH + 4-CHO	32	34	31	40	37	4
OH + 2-$COCH_3$	261	312	326	349	365	374
OH + 3-$COCH_3$	75	114	88	100	98	114
OH + 4-$COCH_3$	27	51	41	57	66	79
OH + 2-$CONH_2$	262	315	371	354	340	281
OH + 4-$CONH_2$	53	100	111	113	(96)	(-8)
OH + 2-CN	47	57	62	70	49	-5
OH + 3-CN	110	115	119	120	117	138
OH + 4-CN	-138	-102	-101	-96	-69	48
OH + 2-OCH_3	32	47	53	69	75	87
OH + 3-OCH_3	22	23	24	29	18	2
OH + 4-OCH_3	-14	-11	-8	-2	-6	-26
OH + 2-CO_2CH_3	253	294	314	340	354	367
OH + 3-CO_2CH_3	60	57	55	47	46	62
OH + 4-CO_2CH_3	30	38	41	44	35	35
OH + 2-CH_3	-10	-8	-10	-15	-18	-27
OH + 3-CH_3	-26	-26	-30	-38	-42	-50
OH + 4-CH_3	-24	-26	-30	-37	-37	-41
OH + 2-NO_2	124	148	181	216	223	128
OH + 3-NO_2	97	109	114	115	116	110
OH + 4-NO_2	12	5	9	25	-25	-118
OH + 2-Ph	-5	18	12	13	-1	-25
OH + 3-Ph	-4	-5	-13	-15	-31	-50
OH + 4-Ph	-5	-8	-17	-24	-32	-56

that interaction index factors can be derived for individual substituents. The interactions of substituents are mutual i.e. each substituent will influence the contribution of each other substituent. The overall term would need to account for the changes in polarity and the modifying effects of substituents on each other. Before any approach of this kind can be followed the sources of interactions must be recognised. The major effects were expected to be due to hydrogen bonding, electron donating-accepting effects and dipole formation (and its extreme case ionisation).

The situation with aromatic substituents is complicated by the possibility of cross ring interactions which would not be expected to occur in aliphatic disubstituted compounds.

For some substituent pairs (e.g. CH_3 + $COCH_3$ in methanol) the RI_{SUM} was within the anticipated error margin (+/- 10) of the experimental retention index (Table 7.12) suggesting that no interactions were occurring. In addition for several substituent pairs there was a close match except at a single value, usually at 80% organic modifier concentration (e.g. CH_3 + 3- and 4-CO_2CH_3). At this modifier composition the capacity factors would be very small and a larger uncertainty in the results would be expected (see Chapter 3). The degree of the interaction depended on the organic modifier and pairs of substituents might show no significant interactions in methanol but larger interactions in acetonitrile.

The substituent pairs in eluents containing methanol (Tables 7.12 and 7.14) for which the interaction increment was less than +/- 10 units (or less than +/- 20 units in 80% methanol) were:-

CH_3 + 2-,3-,4-$COCH_3$	CH_3 + 3-OCH3
CH_3 + 3-,4-CO_2CH_3	CH_3 + 3-NO_2
CH_3 + 3-Ph	CH_3 + 2-,3-CHO
CH_3 + NH_2	CH_3 + 2-,3-,4-CN
OH + 2-CH_3.	OH + 2-Ph

In eluents containing acetonitrile fewer substituent pairs (Table 7.13 and 7.15) had interaction terms which were less than +/-10 units across the eluent range:-

CH_3 + 4-Br	CH_3 + 3-,4-Cl
CH_3 + 3-,4-CH_3	CH_3 + 3-NO_2
CH_3 + 4-Ph	CH_3 + 3-CN

The interaction increments for the other substituent pairs varied considerably in size from -200 to +400 units depending on the substituents and eluents. The largest interaction increments were found for compounds in which intramolecular hydrogen bonding would be expected (retention indices typically increased by between 200 and 400 units). The smallest interaction increments were found for the substituents in which the fixed substituent was the methyl group.

7.2.1 Interaction Indices

The change in the interaction increments across the eluent range could be described by quadratic regression equations to give interaction index equations (Table 7.16

and 7.19). This is probably not the approach which will have to be taken in the long term however due to the amount of data currently available this approach has been adopted as an interim measure to enable the prediction of retention indices. If the change in interaction increment across the eluent range was less than 10 a mean value has been used (a and b = 0). Where there was only a small difference between the isomers (e.g. bromotoluenes and chlorotoluenes) a single common interaction index equation has been derived for the isomers. For the substituent pairs, listed above, with interaction increments < 10 it was assumed that no interactions were occurring and the coefficients of the interaction index equations were set to 0.

7.2.2 Interactions in Disubstituted Compounds

In an analysis of the interactions a number of factors can be considered. It should be noted that it was not possible to fully separate the electronic and other effects such as hydrogen bonding. The observed interaction increments were therefore a combination of the electronic and other effects. However, some of the dominant interactions can however be identified and these can be examined and will be discussed. Although the intra-molecular hydrogen bonding interactions will be discussed as though they were the only interactions occurring in these compounds this is unlikely. A more accurate situation would be to regard the interaction has being in addition to electronic factors in an *ortho* position. Further study may enable a hydrogen bonding term to be calculated in addition to the electronic term , possibly by assuming that the size of the

Table 7.16: Regression equations relating change in interaction increment to methanol concentration for interactions with methyl substituents

$II = ax^2 + bx + c$

x = % modifier

Substituent Pairs	Coefficients of regression equation		
	a	b	c
CH_3 + 2-Br)			
CH_3 + 3-Br)			
CH_3 + 4-Br)	0.0150	-1.230	32
CH_3 + 2-Cl)			
CH_3 + 3-Cl)			
CH_3 + 4-Cl)			
CH_3 + 2-CH_3	0.0221	-1.967	31
CH_3 + 3-CH_3)	0.0207	-1.796	45
CH_3 + 4-CH_3)			
CH_3 + 2-$COCH_3$	0	0	0
CH_3 + 3-$COCH_3$	0	0	0
CH_3 + 4-$COCH_3$	0	0	0
CH_3 + 2-OCH_3	0.0057	-0.306	23
CH_3 + 3-OCH_3	0	0	0
CH_3 + 2-CO_2CH_3	0	0	-11
CH_3 + 3-CO_2CH_3	0	0	0
CH_3 + 4-CO_2CH_3	0	0	0
CH_3 + 2-NO_2	0	0	-28
CH_3 + 3-NO_2	0.0036	-0.119	-2
CH_3 + 4-NO_2	0	0	-12
CH_3 + 2-Ph	0	0	-27
CH_3 + 3-Ph	0	0	0
CH_3 + 4-Ph	0	0	21
CH_3 + 2-CHO	0	0	0
CH_3 + 3-CHO	0	0	0
CH_3 + 4-CHO	0	0	0
CH_3 + 2-$CONH_2$	-0.0429	5.203	-208
CH_3 + 3-$CONH_2$)	0	0	-13
CH_3 + 4-$CONH_2$)			
CH_3 + 2-NH_2)			
CH_3 + 3-NH_2)	0	0	-13
CH_3 + 4-NH_2)			
CH_3 + 2-CN	0	0	0
CH_3 + 3-CN	0	0	0
CH_3 + 4-CN	0	0	0

Table 7.17: Regression equations relating change in interaction increment to acetonitrile concentration for interactions with methyl substituents

$$II = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Substituent Pairs	Coefficients of regression equation		
	a	b	c
CH_3 + 2-Br)			
CH_3 + 3-Br)	0.0232	-2.182	50
CH_3 + 4-Br)			
CH_3 + 2-Cl)			
CH_3 + 3-Cl)	0.0180	-1.875	48
CH_3 + 4-Cl)			
CH_3 + 2-CH_3	0	0	-22
CH_3 + 3-CH_3	0	0	0
CH_3 + 4-CH_3	0	0	0
CH_3 + 2-$COCH_3$	0	0	-11
CH_3 + 3-$COCH_3$)	0	0	-19
CH_3 + 4-$COCH_3$)			
CH_3 + 2-OCH_3	0.0062	-0.325	18
CH_3 + 3-OCH_3	0	0	-15
CH_3 + 2-CO_2CH_3	0	0	-16
CH3 + 3-CO_2CH_3)	0	0	-10
CH_3 + 4-CO_2CH_3)			
CH_3 + 2-NO_2	0	0	-25
CH_3 + 3-NO_2	0	0	0
CH_3 + 4-NO_2	0	0	-12
CH_3 + 2-Ph	0	0	-21
CH_3 + 3-Ph	0	0	0
CH_3 + 4-Ph	0	0	0
CH_3 + 2-CHO	0.0073	-0.454	-7
CH_3 + 3-CHO	-0.0025	0.400	-14
CH_3 + 4-CHO	-0.0075	-0.428	-16
CH_3 + 2-$CONH_2$	0	0	-58
CH_3 + 3-$CONH_2$)	-0.0288	2.631	-82
CH_3 + 4-$CONH_2$	0	0	-29
CH_3 + 2-NH_2)			
CH_3 + 3-NH_2)	0	0	-27
CH_3 + 4-NH_2)			
CH_3 + 2-CN	0	0	-20
CH_3 + 3-CN	0	0	0
CH_3 + 4-CN	0	0	-15

Table 7.18: Regression equations relating change in interaction increment to methanol concentration for interactions with hydroxyl substituents

$$II = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Substituent Pairs	Coefficients of regression equation		
	a	b	c
OH + 2-NH_2	0.0471	-2.577	161
OH + 3-NH_2	0.0543	-3.274	78
OH + 4-NH_2	0.1707	-15.756	363
OH + 2-Br	-0.0550	5.790	-155
OH + 2-Cl	-0.0343	3.574	-96
OH + 3-Br)			
OH + 4-Br)	-0.0170	1.777	-7
OH + 3-Cl)			
OH + 4-Cl)			
OH + 2-OH	0.0736	-5.439	216
OH + 3-OH	0.0621	-4.787	142
OH + 4-OH	0.0779	-5.493	87
OH + 2-CHO	0.0129	1.017	154
OH + 3-CHO	-0.0021	1.067	49
OH + 4-CHO	0	-2.150	122
OH + 2-$COCH_3$)	0.0321	0.253	193
OH + 2-CO_2CH_3)			
OH + 3-$COCH_3$	0.0420	-3.627	159
OH + 4-$COCH_3$	-0.1240	12.914	-274
OH + 2-$CONH_2$	-0.0214	3.551	147
OH + 4-$CONH_2$	0	1.550	8
OH + 2-CN	-0.1230	10.445	-161
OH + 3-CN	0.0150	-0.730	125
OH + 4-CN	-0.0725	7.205	-97
OH + 2-OCH_3	0.0079	-0.313	36
OH + 3-OCH_3	0	0	18
OH + 4-OCH_3	0	0	-16
OH + 3-CO_2CH_3	-0.0179	2.493	-11
OH + 4-CO_2CH_3	-0.1057	11.186	-231
OH + 2-CH_3	0	0	0
OH + 3-CH_3)	0.0114	-1.811	51
OH + 4-CH_3)			
OH + 2-NO_2	0	-2.000	200
OH + 3-NO_2	0	0.210	96
OH + 4-NO_2	0	-3.800	145
OH + 2-Ph	0	0	-9
OH + 3-Ph)	0.0076	-1.313	46
OH + 4-Ph)			

Table 7.19: Regression equations relating change in interaction increment to acetonitrile concentration for interactions with hydroxyl substituents

$II = ax^2 + bx + c$ $\quad$ x = % modifier

Substituent Pairs	Coefficients of regression equation		
	a	b	c
OH + 2-NH_2	0.0511	-1.124	94
OH + 3-NH_2	0.0388	-2.525	88
OH + 4-NH_2	0.0759	-5.794	103
OH + 2-Br	-0.0134	0.670	-12
OH + 2-Cl	-0.0018	-0.798	66
OH + 3-Br)			
OH + 4-Br)	-0.0052	-0.490	53
OH + 3-Cl)			
OH + 4-Cl)			
OH + 2-OH	-0.0386	6.737	-72
OH + 3-OH	0.0145	0.620	20
OH + 4-OH	0.0246	0.455	-36
OH + 2-CHO	0.0454	-3.000	355
OH + 3-CHO	-0.0100	1.380	48
OH + 4-CHO	-0.0313	3.089	-37
OH + 2-$COCH_3$)	-0.0311	5.598	121
OH + 2-CO_2CH_3)			
OH + 3-$COCH_3$	0.0059	0.128	65
OH + 4-$COCH_3$	0.0071	0.171	20
OH + 2-CONH2	-0.1500	16.937	-114
OH + 4-$CONH_2$	0	1.910	8
OH + 2-CN	-0.0757	7.540	-117
OH + 3-CN	0.0092	-0.604	122
OH + 4-CN	0.0909	-7.044	10
OH + 2-OCH_3	-0.0027	1.366	-6
OH + 3-OCH_3	-0.0100	1.380	48
OH + 4-OCH_3	-0.0309	3.373	-95
OH + 3-CO_2CH_3	0.0177	-2.033	108
OH + 4-CO_2CH_3	-0.0157	1.783	-9
OH + 2-NO_2	-0.0486	7.877	-88
OH + 3-NO_2	-0.0350	4.356	-16
OH + 3-NO_2	-0.1314	12.646	-278
OH + 2-CH3	-0.0105	0.816	-25
OH + 3-CH_3)	-0.0030	-0.109	-18
OH + 4-CH_3)			
OH + 2-Ph	-0.0477	4.799	-104
OH + 3-Ph)	-0.0198	1.266	-26
OH + 4-Ph)			

ortho electronic interactions was similar to that of the *para* isomer although the validity of such an approach might be difficult to assess.

a) Interactions in Hydroxyl and Carbonyl Containing Compounds

Four compounds were examined in which there were *ortho* hydroxyl and carbonyl groups (OH + CHO, OH + CO_2CH_3, OH + $COCH_3$, and OH + $CONH_2$). These compounds can undergo strong intra-molecular hydrogen bonding and in all these compounds the interaction can be clearly seen in the much higher interaction increments for the *ortho* isomer when compared with the *meta* and *para* isomers. In each case an increase in retention index, relative to the RI_{SUM}, of between 200 and 450 retention index units (Figure 7.1, Tables 7.14 and 7.15), was found. Although a single interaction index expression may not be sufficient to account for all these interactions a common term has been used for *ortho* OH + CO_2R and OH + COR, it should be remembered that electronic interactions also occur in addition to hydrogen bonding which may account for differences between the different substituent pairings.

b) Interactions in Hydroxyl - Hydroxyl and Hydroxyl - Amino Compounds

Hydrogen bonding can also occur between OH + OH and OH + NH_2. In these compounds the difference in retention index between the different isomers was not as large as with the carbonyl containing compounds. For OH + 2-OH and OH + 2-NH_2

the interactions caused an increase in retention index of between 100 and 250 units compared with 200 - 450 units in carbonyl containing compounds. The change in the increment across the eluent range was greater than the corresponding change with hydrogen bonding carbonyl groups which may reflect the influence of electronic effects.

To a first approximation it might probably be possible to use a single interaction index equation for the hydrogen bonding *ortho* amino and hydroxyl substituents. However, this was not done as in both pairs of substituents the *meta* and *para* isomers had fairly large and significant interaction increments which were different for the different positions showing that these substituents were susceptible to electronic interactions.

c) Interactions in Hydroxyl - Nitro compounds

In *ortho* substituted nitrophenol intramolecular hydrogen bonding may occur[177] but this would be expected to be balanced by electronic effects and also intermolecular hydrogen bonding with the eluent. In acetonitrile the interaction of the *ortho* nitro containing compound was of a similar magnitude to that observed for the the *ortho* amino and hydroxyl compounds suggesting that intramolecular hydrogen bonding may be occurring, however, the interaction in methanol was smaller and at high methanol concentrations the compound was unretained on the column.

In both organic modifiers the *meta* isomer also showed a large interaction but the *para* isomer had a much smaller interaction. The possibility of ionisation of the *para* isomer was discussed previously and may be reflected in the

negative interaction increments observed for the *para* isomer.

d) Interactions in Compounds Without Hydrogen Bonding Groups

The other interactions which could occur between the *ortho* substituents were electronic interactions and steric interactions. The two effects would probably be superimposed and the resulting interaction increments would be the sum of the two interactions.

If the fixed substituent was a methyl group (Table 7.12 and 7.13) the *ortho* interaction increment was frequently found to differ from the other two isomers, which usually had similar interaction indices. This was also found by Morishita[114] studying group contribution effects for the pairs CH_3 + CH_3 and CH_3 + OH. The major exception to this effect was the methyl + halides and amines in which the three isomers had very similar interaction increments, although these were small and considerably less than the electronic effects observed for other substituent pairs. In most cases there was a reduction in the retention index of the *ortho* isomers when compared to the other isomers although the size of the interaction (-70 to +7 units) depended on the substituent pairs and the eluent composition. The largest effects were observed with the bulkier amide, nitro and ester group in which steric effects would reduce co-planarity with the ring.

The *meta* and *para* isomers with a methyl substituent had similar interaction indices with most substituent pairs. As the methyl group would not be expected to contribute significantly to any electronic interactions this would be

expected again. The size of the interaction varied considerably (-30 to +20 units), in most cases the value was almost constant across the eluent range.

If the fixed substituent was a hydroxyl group and hydrogen bonding interactions would not be expected the interaction increments (0 - 50 units, Tables 7.14 and 7.15) were considerably smaller than the hydrogen bonding interactions. The interactions between *ortho* substituents other than hydrogen bonding would be expected to be a combination of electronic interactions between the substituents and proximity effects.

For substituents *meta* or *para* to each other electronic interactions would be expected to occur these would depend on both the nature of the two substituents (the electron donating / accepting strength) and the positions. Generally the smallest interaction increments (within a single set of compounds i.e. with a fixed hydroxyl group) occurred when the two substituents were electron donating e.g. hydroxy + methoxy. However, hydroxy-hydroxyl and hydroxyl-amine pairings were exceptions and in most cases there were large changes with eluent composition. In most cases the absolute value of the interactions was higher than the methyl *meta* and *para* isomers, probably due to the interactions involving the hydroxyl group. Consequently it is not possible to make any generalisations about the groupings which produce the largest interactions.

7.3 RELATIONSHIP BETWEEN INTERACTION INCREMENTS AND STUDIES ON OCTANOL-WATER PARTITION COEFFICIENTS

The substituent indices calculated for individual aromatic and aliphatic substituents were found to be linearly related to the Hansch substituent contributions (π) to the octanol-water partition coefficients (Chapter 4 and 6). The compilation of π values[48] mainly contains values derived for single substituents on benzene. However, a simple summation of the π terms to calculate log *P* was not successful for compounds in which interactions between substituents occurred. Initially this led to the development of different sets of π values[178], for example with aniline or phenol, as the parent. This approach was clearly unsatisfactory, as in the present study (see earlier), because it involved a large proliferation of data. Alternative approaches have only been made relatively recently49-52. Initially the methods for log *P* "correction" ignored *ortho* substituents as it was recognised that these compounds could show complex interactions and these have only been included very recently. The aims of Fujita[49,50] and Leo[51,52] have been to develop methods to calculate log *P* values for poly-substituted compounds using a set of ground rules to account for the interactions. As many interactions in the octanol-water and RP-HPLC system would be expected to be similar, this approach will be discussed in some detail to see whether it would be possible to apply a similar method to the current interaction increments in order to replace what would be a long list of empirical values by generally applicable parameters.

Fujita[49,50] attempted to describe the substituent

interactions in octanol-water partition coefficients using Hammett-type relationships. Initially he[49] examined only *meta* and *para* isomers. He found that the difference between the π value derived with benzene as the parent and that with phenol as the parent could be described using the equation

$$\delta\pi = \pi_{X(PhOH)} - \pi_{X(PhH)} = 0.82\sigma_X + 0.06 \qquad 7.2$$

in which σ was the Hammett constant of the substituent X. Effectively the Hammett constant was being used as a measure of the change in the polarity of the molecule relative to the mono-substituted compounds. Several compilations of Hammett constants exist so their use may be possible. Although equations of the type shown above were found to hold true for systems in which the parent was phenylacetic acid, phenoxyacetic acids or benzyl alcohol, it could not be used for systems in which the parent was either benzoic acid or nitrobenzene. Fujita[49] reported that this was due to a "change" in the relative importance of the two substituents in the molecule and that in the latter two systems the acid group or nitro group dominated the interactions and therefore the determination of log *P*. This led to a more general equation which attempted to "quantify" the effect of one group on another across the aromatic ring in terms of a "susceptibility" constant (*p*) and Equation 7.3:-

$$\pi_{X(PhXY)} - \pi_{X(PhH)} = p_Y\sigma_X + p_X\sigma_Y \qquad 7.3$$

p_X and p_Y are the susceptibilities of X and Y to the modifying effects of Y and X, respectively. Separate Hammett constants (σ) were used for the *meta* and *para* isomers, *ortho*

isomers were not considered. The observed changes in the octanol-water partition coefficient were explained in terms of the the effect of the substituent on the ability of the compound to hydrogen bond with the solvent. Consequently the *p* term for a non-hydrogen bonding substituent (CH_3, H, Halogen, Ph, CF_3 and SF_5) was zero, although the Hammett constant was not zero. With these substituents as the Y substituent the equation became:

$$\pi_{X(PhXY)} - \pi_{X(PhH)} = p_X\sigma_Y \qquad 7.4$$

The other term dominated when the substituent X hydrogen bonded with the solvent. If neither substituent hydrogen bonded with the solvent the interaction term would be negligible. The susceptibility constants were derived experimentally using multiple regression analysis on a set of compounds with fixed substituents and variable substituents.

In this approach *ortho* substituents were not initially considered. However, the approach has now been extended to *ortho* substituents although these are treated separately from *meta* and *para* substituents. The *ortho* substituent effects were divided into two terms an electronic effect equivalent to that in *para* isomers and an additional term to account for the proximity effects (Equation 7.5):

$$\pi_{o-XPhY} = a\pi_{XPhY} + p_Y\sigma^{\sigma}_{o-X} + p_X\sigma^{\sigma}_{o-Y} + f_YF_X + f_XF_Y \qquad 7.5$$
$$+ \delta E^{X}_{m} + \delta E^{Y}_{m} + c$$

Where $a\pi$ = the additive π, E = Taft steric effect value, F = Swain-Lipton field effect constant, f = the susceptibility

for the *F* constant and δ the susceptibility for the Taft *E* value. The *E* term was found to be insignificant for phenols and anilines (in the absence of intramolecular hydrogen bonding). The susceptibility constants were derived from a limited set of data.

Leo[51,52] has also examined the interactions which occurred between substituents and their effect on octanol-water partition coefficients. He[51,52] recognised that the interactions were not purely electronic but also involved other factors such as hydrogen bonding, steric interactions, interactions in mixed alkyl-aryl compounds and the electronic interactions. He attempted to simplify the approach to electronic interactions taken by Fujita to facilitate rapid estimation of the interaction terms. The electronic effects were calculated using a similar approach to that of Fujita (described above). The contribution was calculated as the product of the susceptibility and Hammett type constants In this case the constant was not the actual Hammett constant but a "hydrophobic" σ/p value calculated from a set of model compounds with well defined interactions. Several assumptions were made to obtain the electronic interaction term, firstly the value was assumed to be the same for *meta*, *ortho*, and *para* isomers, and a single σ and *p* value was used for each member of a class of substituents. The substituents were divided into three classes either inducers with $p = 0$ (e.g. CN, Br), responders with $\sigma = 0$ (e.g. OH, NH_2) or bi-directional with *p* and $\sigma > 0$ (e.g. CHO, CO_2CH_3). For the bi-directional substituents the overall effect was governed by the second substituent present in the compound. In addition to the *ortho* effects of hydrogen bonding a "negative" *ortho* effect was also observed

with some substituent pairs, this was observed to occur with hydroxyl substituted compounds but not amino substituted.

The octanol water partition coefficient was then calculated from the sum of the components of the individual interactions. Multiple regression equations were again used to examine the significance of the different parameters and an overall equation was derived[32].

$$\log P = ALP + F_{\sigma} - 0.29F_{o} + 0.63F_{HB} - 0.15F_{\alpha} \qquad 7.6$$

The terms were ALP = additive log *P*, electronic interactions (F_{σ}), *ortho* effects (F_{o}), hydrogen bonding (F_{HB}) and alkyl-aryl effects (F_{α}). With the exception of the electronic effects each of the terms was quantized taking values 0, 1, 2 etc. The *ortho* effect was considered to be in addition to the other electronic effects and was not present in all *ortho* substituents. Its presence was determined experimentally in the set of data used to calculate the interactions. The electronic term was the same for all isomers of a compound and was initially calculated as the mean of the *meta* and *para* substituent terms. Although a single intra-molecular hydrogen bonding term was suggested this was found to be insufficient to account for the observed hydrogen bonding effects of *ortho* hydroxyl and amide groups. The hydrogen bonding term was not used for substituent pairs such as *ortho* OH + OH and *ortho* OH + NH_2 but the interaction terms were assumed to be equal to that of the *meta* and *para* substituents. The final term in the equation (F_{α}) was used to account for differences between the additive log *P* and the experimental values for compounds with alkyl groups. These included compounds such as

dimethylbenzenes. In these compounds the multiplier was calculated as 1 less than the number of methyl groups (up to a maximum of 3).

7.3.1 Possible Application of the Above Approach to the Current Study

As many Hammett constants are available[48,179] and also the derived susceptibility constants have been listed by Fujita[50] and Leo[52] an attempt was made to determine whether this approach would be applicable to the current work. The interactions between the substituents would be expected to be similar to those occurring in the octanol-water partition coefficient although significant differences might be expected for some substituents. For example the hydroxyl group has been shown to be an outlier in the substituent index-π relationship and the example interactions used here may therefore not be the ideal test as to whether the values could be applied directly.

a) Relationship between Interaction Increments and Hammett Constants for Hydroxyl Substituted Compounds

Initially a study was carried out using the Hammett constants for phenol substituted compounds, these were listed by Barlin and Perrin[179]. This method was similar to the approach first used by Fujita *et al.*[49]. Apparent values of the Hammett constants were also listed for *ortho* substituents on phenols and these were used for the *ortho* substituted compounds. These Hammett constants do not measure the "interaction" as such but were the difference

Figure 7.2: Relationship between apparent Hammett constant for phenol substituted compounds and substituent increment

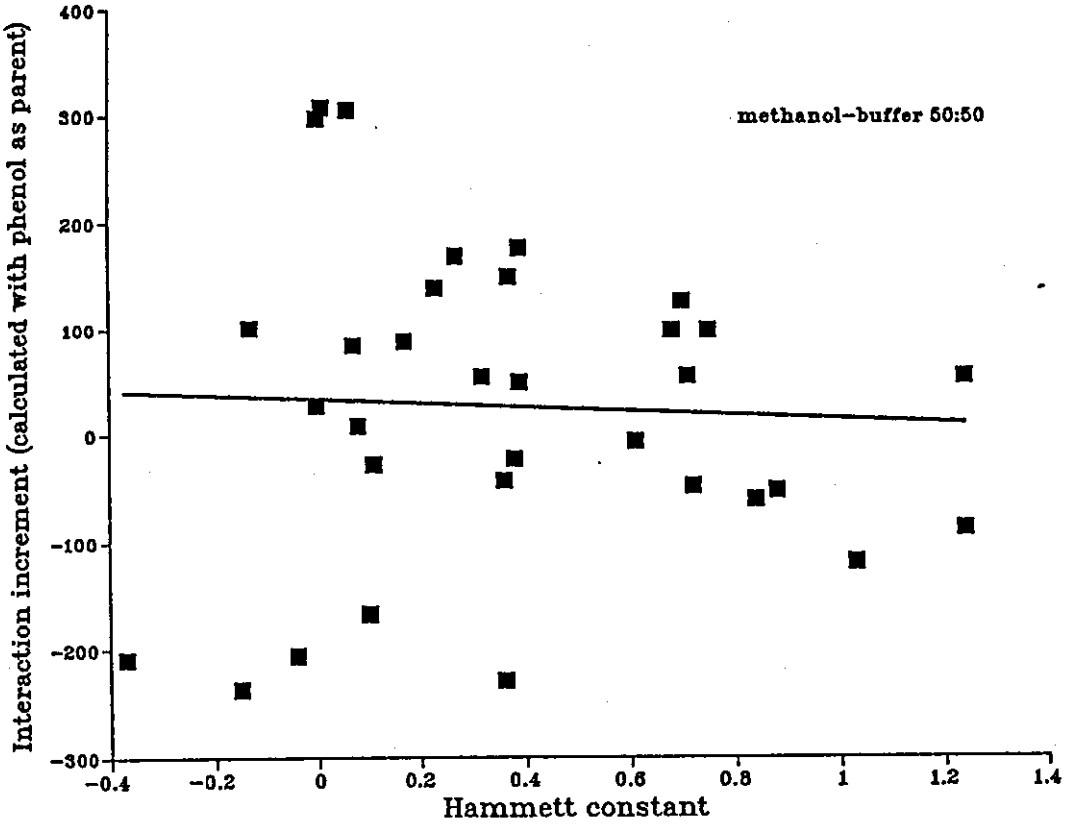
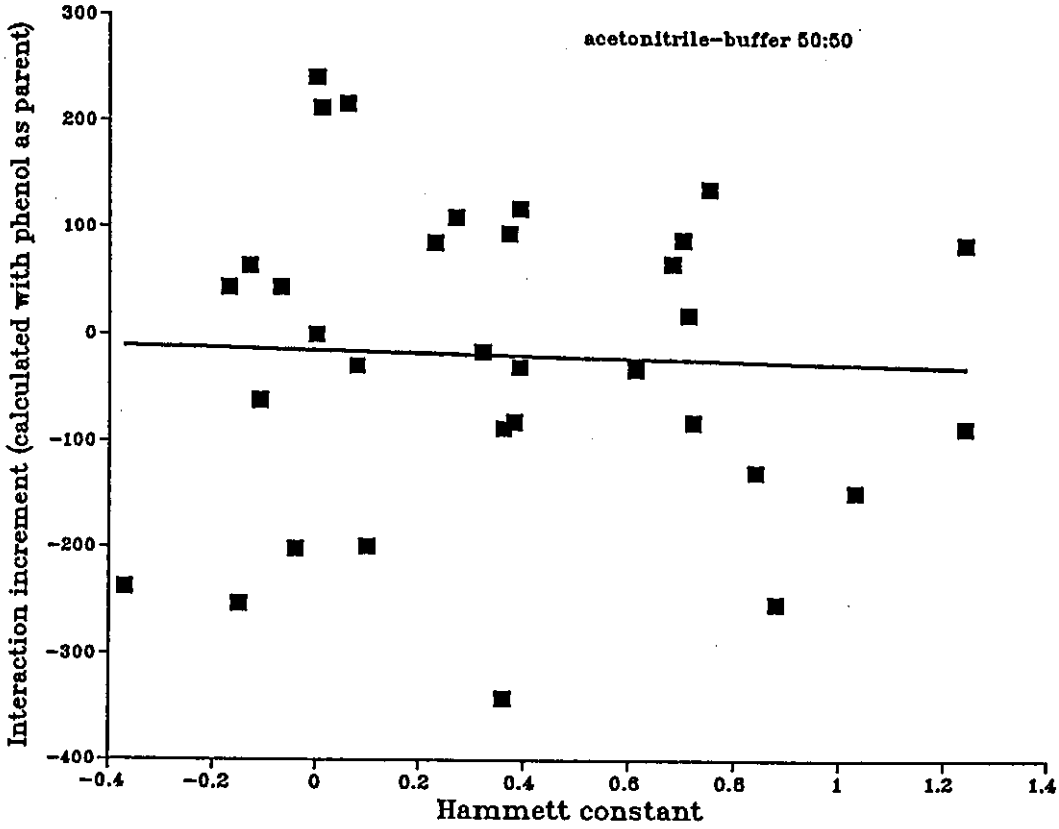


Figure 7.3: Relationship between apparent Hammett constant for phenol substituted compounds and substituent increment



between the measured values and that of phenol. They would therefore be expected to be related not to the interaction increments tabulated in this chapter (Tables 7.16 to 7.19) but to increments calculated using phenol as the parent. These values were therefore calculated as :-

$$\delta I_{OH} = RI_{XPhOH} - RI_{PhOH} \qquad 7.7$$

However, no relationship between the Hammett constants and substituent effects was observed (Figure 7.2 in 50% methanol and Figure 7.3 in 50% acetonitrile). On their own these values are not sufficient to account for any interactions. A simple direct application of the Hammett constants does not therefore appear appropriate.

b) Relationship between Leo method of correcting for interaction and interaction increments

Leo[52] listed a number of susceptibility values and "hydrophobicity" σ values. A brief study was carried out to see whether the interaction terms calculated by Leo's equation (below) could be correlated with the interaction increments calculated in this study. The interaction terms for the substituent terms were calculated as

$$IT = p_X\sigma_Y + p_Y\sigma_X - 0.29F_O + 0.63F_{HB} - 0.15F_\alpha \qquad 7.8$$

Using the values suggested by Leo[51,52] (Table 7.20) the interaction terms for the substituent pairs were calculated (Table 7.21). The interaction terms would be expected to be related to the interaction increments obtained in this

Table 7.20: Terms suggested by Leo to calculate interactions (taken from References 52 and 51)

Substituent	σ	p	F_{HB}	F_{σ} OH	F_{σ} CH_3	$F_{\alpha\cdot}$
CN	0.65	0	0	0	0	0
NO_2	0.60	0	0	0	0	0
Br	0.28	0	0	1	0	0
Cl	0.28	0	0	1	0	0
CHO	0.58	0.44	1	-	0	0
CO_2CH_3	0.51	0.27	1	-	0	0
$COCH_3$	0.51	0.27	1	-	0	0
$CONH_2$	0.32	0.72	1	-	2	0
OCH_3	0.17	0.50	0	1	0	0
OH	0	1.06	0	0	0	0
NH_2	0	1.08	0	0	0	0
CH_3	0	0	0	0	0	1

Table 7.21: Interaction terms calculated using equation 7.8 and the values listed in Table 7.20.

$$IT = p_X\sigma_Y + p_Y\sigma_X - 0.29F_\sigma + 0.63F_{HB} - 0.15F_\alpha.$$

(Compound XPhY)

Y	Position	IT X = OH	IT X = CH_3
CN	2	0.689	0
CN	3	0.689	0
CN	4	0.689	0
NO_2	2	0.636	0
NO_2	3	0.636	0
NO_2	4	0.636	0
Br	2	0.007	0
Br	3	0.297	0
Br	4	0.297	0
Cl	2	0.007	0
Cl	3	0.297	0
Cl	4	0.297	0
CHO	2	1.020	0
CHO	3	0.573	0
CHO	4	0.573	0
CO_2CH_3	2	1.171	0
CO_2CH_3	3	0.541	0
CO_2CH_3	4	0.541	0
$COCH_3$	2	1.171	0
$COCH_3$	3	0.541	0
$COCH_3$	4	0.541	0
$CONH_2$	2	0.969	-0.241
$CONH_2$	3	0.339	0
$CONH_2$	4	0.339	0
OCH_3	2	-0.110	0
OCH_3	3	0.180	0
OCH_3	4	0.180	0
OH	2	0	0
OH	3	0	0
OH	4	0	0
NH_2	2	0	0
NH_2	3	0	0
NH_2	4	0	0
CH_3	2	0	-0.15
CH_3	3	0	0
CH_3	4	0	0

study, therefore the interaction increments and interaction terms have been plotted in the two eluents methanol-buffer 50:50 and acetonitrile-buffer 50:50 (Figure 7.4 and 7.5). In this case there appears to be a general trend although the correlation would not be high. One of the main areas where the Leo model failed to account for the interaction was with the methyl substituents for which no interactions terms could be calculated but differences between the experimental *RI* and RI_{calcm} were observed. The assumption of equal electronic interactions for all three positions could also present problems. The terms derived by Leo could not be directly applied to the current system although there is the possibility of deriving similar interaction terms using HPLC systems. The values used by Leo were based on 400 - 500 substituted compounds so in the present study insufficient interaction increments may have been collected to derive reliable values.

7.4 APPLICATION OF INTERACTION INDICES

As only a limited number of compounds have been studied it was not possible to do an analysis similar to that of Leo (described above) for the HPLC data. It is therefore necessary to include the individual empirical interaction index equations to predict the retention indices of unknown compounds. This approach is not satisfactory and has only been adopted as an interim measure until an improved method of accounting for interactions can be derived.

Figure 7.4: Relationship between interaction increment and Hansch interaction terms

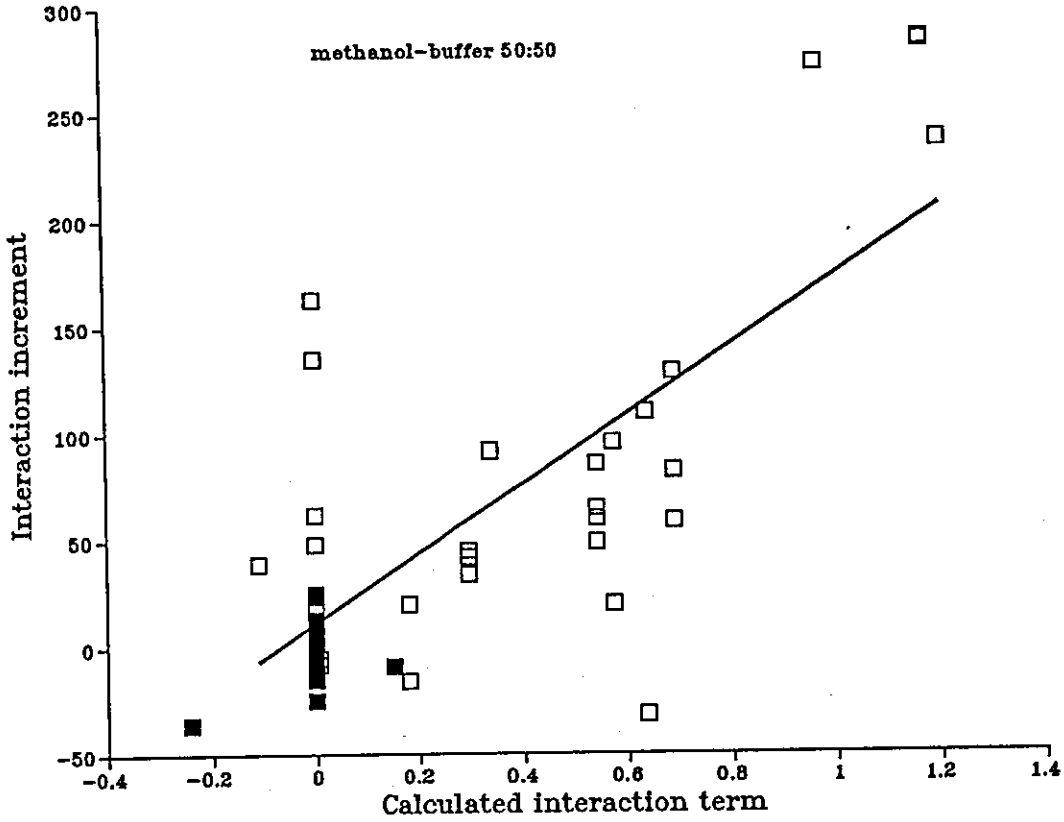
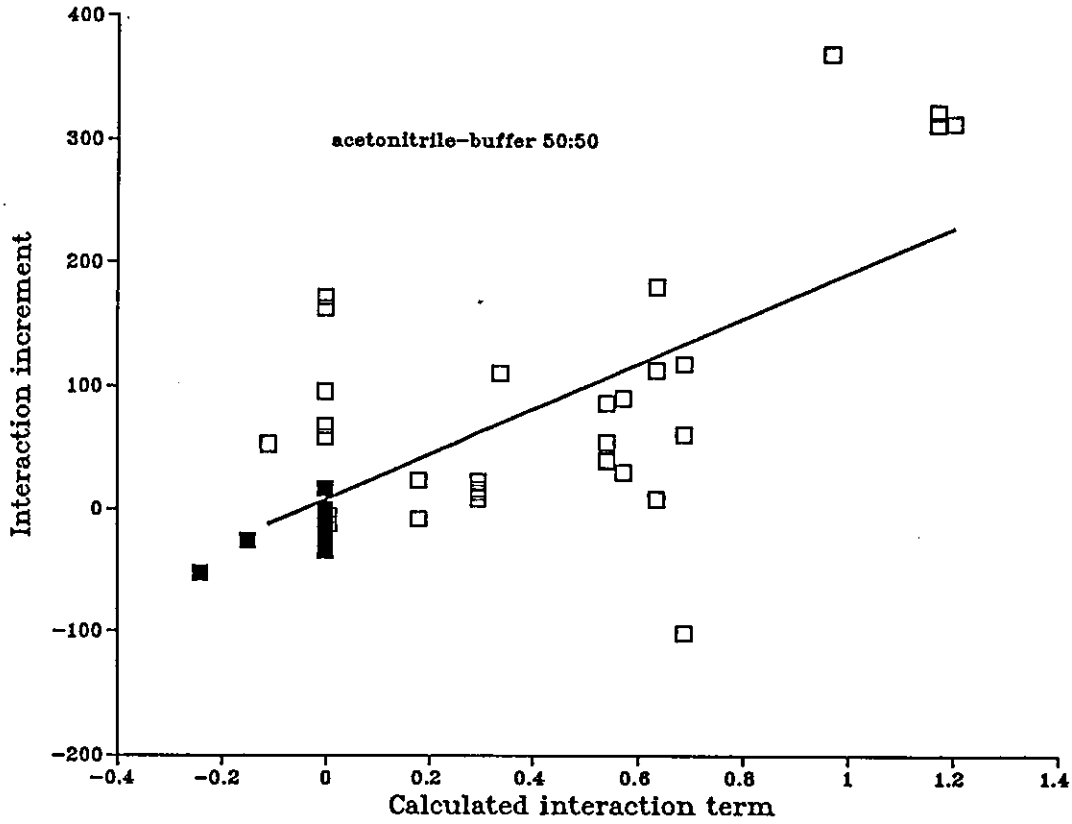


Figure 7.5: Relationship between interaction increment and Hansch interaction terms



CHAPTER 8

IMPLEMENTATION OF THE PREDICTION METHOD USING AN EXPERT SYSTEM

In the preceding chapters the collection of the data required to predict retention indices has been discussed. The changes in the parameters, with eluent composition, have been described using quadratic regression equations. These have been collected together to into a set of spreadsheets which form the basis of the retention prediction system. The next stage in developing the prediction system was to devise a method by which the data can be easily recalled from the spreadsheet. Although it would be perfectly possible to do this manually using pen, paper and a calculator this task would be extremely time consuming and inefficient because it would be easy to miss some interactions which should be included. An expert system program, Chromatographic Retention Index Expert System, CRIPES, has therefore been used to provide a "user-friendly" interface with the spreadsheet. In this chapter a brief description of how an expert system works and a summary of its use in chromatography will be given. Following this the expert system used for this work will be discussed together with a description of the testing of the system.

8.1 USE OF COMPUTER PREDICTION METHODS IN HPLC

Attempts to automate the development of separation methods have led to the use of various automatic

optimisation packages. In these the program attempts to optimise a predefined criteria, for example the resolution or chromatographic response factor[180-184]. In most cases the programs require a considerable input of experimental data at a minimum of two eluent compositions or in two organic modifiers. Frequently the programs work on a "trial and error" or iterative approach with the user performing separations at the proposed optimum conditions and then inputting the information back into the program to search for a new optimum. Although the process can be automated the amount of chromatographic data required can be quite large.

Alternatively there have been methods which use the retention based on a gradient elution profile to estimate isocratic retention of the analytes[169]. None of these methods can be said to be true prediction but rather extrapolation or interpolation between experimental points. Frequently an assumption of a linear relationship between the retention (log capacity factor) has been used as the basis of the systems. As was shown earlier (Chapter 4, Section 4.2) this assumption was not valid for all compounds particularly over a large eluent range. An automated system, Drylab, based purely on the linear relationship between log k' and the eluent composition has been proposed by Snyder. This has been investigated briefly and a description will be given in a later section.

None of the optimisation systems of the types described above make any use of the structure of the analytes in determining the optimum separation conditions. A method proposed by Jinno *et al.*[29,30,77,78] attempted to use structural parameters to predict retention and to optimise

the separation. Regression equations relating different compound types to structural parameters were held in a spreadsheet. The structural parameters used depended on the type of compound and multiple regression analysis was used to obtain the equations for the individual column being used. A computer program was then used to extract data from the spreadsheet. The user was required to input information on the compound classes e.g. phenol, alkylbenzene or polynuclear aromatic hydrocarbon. Using this data the program would extract the retention parameters and the user would then be required to input structural information on the compounds to be separated. The retention (log k') of the compounds could then be calculated and the separation optimised without any further experimental work. There are several limitations of this work, firstly it requires the user to know several structural parameters for all the solutes (log P parameters, Hammett constants, Van der Waals radius/volume, the number of hydrogen bonding and accepting groups etc.) with different parameters being applied for different types of compounds and on different stationary phases. As the method is based on capacity factors some initial experimental work with reference compounds must be performed on every column to obtain the initial regression equations before any prediction can be carried out.

Several other prediction methods described in Chapter 1 have been developed using computers but none of these has been developed as extensively as that of Jinno and co-workers. A recent development has been a number of proposals which have used computer based expert systems for retention prediction and these will be discussed in a later section of this chapter (Section 8.3).

8.2 INTRODUCTION TO EXPERT OR KNOWLEDGE BASED INFORMATION SYSTEMS

The aim of this section is not to provide a detailed explanation of the workings of an expert system or even to define what is meant by the term expert system. Indeed the term "expert system" is probably a great misnomer but it is considerably easier to use than the alternative "Knowledge Based Information Systems", which would probably be a more accurate description of all the systems discussed in this chapter. The aim is to give a very brief description of how an expert system works and to provide a definition of the various terms which will be used in the chapter. Further information on the workings of expert systems can be found in many books on artificial intelligence including references in this chapter[185,186] .

There is no consensus as to what constitutes an expert system, however, there appear to be several common features. Primarily an expert system is a piece of computer software. The expert system can be divided into two parts the "inference engine" and the "knowledge base". The knowledge base contains the information about a particular well defined "domain" and provides the basis for any consultation using the expert system. The inference engine contains the mechanism by which the computer can extract and interpret data contained within the knowledge base. The form of the knowledge base can differ but the commonest type of expert system is the "rule based" system which consists of a number of "Rules", of the form "If X = Y then W = true", and statements of facts. It has also been suggested that an expert system needs to contain an explanation facility by

which the reasoning behind any answer can be discovered. The complexity of the explanation facility appears to vary greatly between different expert systems.

Within the description of an expert system a few other distinctions need to be made. An expert system may be of two types, either a totally purpose built system designed for a specific application or a program written for a "shell". An expert system "shell" consists of a generally applicable inference engine to which the user adds a knowledge base. An expert system shell provides a flexible approach to expert systems in that the same inference engine can then be used for several completely separate problems by choosing the appropriate knowledge base. It is also often possible to write the knowledge base without specific knowledge of the computer language in which the inference engine is written. The advantage of a purpose written expert system is that it can contain all the features that the user thinks will be required for interpretation of the particular problem. The use of a shell is a compromise which may require the user to rethink an approach to a problem. The program which was used in this work, which will be discussed later, is an example of the use of a shell (VP-Expert). Many similar shells are available in a variety of sizes, degrees of sophistication and price.

Different types of expert systems exist although to date most of the systems used in the analytical chemistry domain are rule-based systems. The programs work by one of two methods, either backward chaining or forward chaining. In the backward chaining mode the program is given a goal which it attempts to achieve by examining the rules. The program identifies each rule which contains the goal as its

conclusion and checks to find if the conditions stated in the rule are satisfied. If the conditions are true then the goal is achieved and the program moves onto its next goal. In a forward chaining system a conclusion is obtained from the facts and rules presented to the system. Many expert systems, including VP-Expert, can to some extent use both methods of chaining.

8.3 USE OF EXPERT SYSTEMS IN CHROMATOGRAPHY

Expert systems have been used in chemistry for many years, one of the earliest applications was the Dendral project started in 1965 which was concerned with the interpretation of mass spectra data[185]. The application of artificial intelligence techniques and expert systems to chemistry, and in particular analytical chemistry has been the subject of several recent articles[187-196]. There appears to be a consensus of opinion that this is an area where expert systems can play an increasingly important role. Chromatography has been identified by several workers as a suitable "domain" for the application of expert systems. There are two distinct approaches to the use of expert systems in HPLC, these are to develop systems that attempt to take into account every possibility and eventuality in the domain or to restrict the application to small well defined problem areas.

An introduction to the use of expert systems in analytical chemistry was presented in a series of tutorials published in Analytical Chemistry in 1984[195,196]. In this article Karnicky explained the aims of the group at Varian in developing ECAT, Expert Chromatographic Assistance Team.

The aim appeared to provide a complete coverage of the field of chromatography choosing between gas and liquid chromatography, the experimental conditions and the detectors. Although descriptions of the work have appeared frequently at conferences (e.g HPLC 87[197]) relatively little has appeared in press[190,196,198]. It is not clear whether there is a real possibility of the program appearing commercially.

Lu Peichang and co-workers have discussed the possibility of developing a chromatograph with "artificial intelligence"[199,200,201]. The expert system appears to act as method of selection of mobile phase conditions and column systems as well as instrumental factors such as detectors and contains information on GC and HPLC separation. No details on the implementation of the expert system are available.

Tischler and Fox[202] described an expert system whose aim was to aid the "inexperienced analytical chemist in choosing a separation method for HPLC". The program, ESP, Expert Separation Program, was a rule based system with the rules being drawn from a standard textbook.

A specific application of an expert system program was discussed by Gunasingham[203]. A program was written to plan the separation of steroids. The user was required to enter details about the sample, such as the polarity of the steroids, the class and the origin of the sample. The program would then detail any sample preparation required and the separation conditions. The same workers have recently described another approach to HPLC optimisation using an expert system. The structure of the program was given but there was no indication of the success of the

method[204].

Bridges *et al.*[205,206] have described the the use of an expert system for eluent optimisation using a diode array detector. The method was similar to automated optimisation systems in that information from a gradient run and isocratic runs was required to determine the separation conditions. The optimisation routine was based on a method of simplex optimisation.

Musch *et al.*[207,208] have written an expert system to decide between UV and amperometric detection for HPLC separation of pharmaceutical compounds. This program was written for an expert system shell, KES. The same workers have suggested how an expert system could be used for automated method development although no details were given[209]. These workers are also involved with the EEC sponsored project which will be discussed below.

An indication of the current interest in expert systems is the EEC sponsored project, ESCA, Expert Systems for Chemical Analysis, being studied under the ESPRIT program by groups in U.K. Netherlands, and Belgium. An overview of the aims of the project was recently published by Schoenmakers and Mulholland[210]. Four possible areas for the application of expert systems to method development for pharmaceutical analysis were identified. These were selection of the initial HPLC conditions, selection of the criteria for selectivity optimisation, optimisation of the chromatographic parameters and method validation. Further details of the approach being taken in each of the application areas were presented at a recent conference[211-215]. Contained within the first application area (the initial selection of chromatographic conditions) is an

attempt to predict the separation conditions based on the presence or absence of structural units in a particular drug molecule[212]. Each area is being implemented using a different expert system shell to compare the applicability of different approaches. The project is not currently seen as a commercial package but more an evaluation of the possible scope for the use of expert systems.

Other expert systems have been described in an article by Glajch[190] but no details of the specific programs were given. The applications described included a program to model the HPLC profile of protein digests (HIPERCALC) by Hodges and a program to simulate peptide separations was also described by Sasagawa. In this article the Drylab program by Snyder *et al.*[167-169] was also described as an expert system. This software does not fit the description of an expert system detailed in the first section of this chapter although it is described as such in the manual[216]. The use of this program will be described later in this chapter.

With the exception of the Drylab software none of the expert systems for optimisation are commercially available. One example of a commercially available expert system for chromatography is HPLC Doctor, LC Resources. This system is aimed at diagnosing malfunctions occurring in HPLC systems but it has not been possible to examine it directly. This area of fault diagnosis has been developed in fields other than chromatography and is one area in which an expert system is useful in enabling a "novice" user to tap the experience of "experts".

8.4 INTRODUCTION TO VP-EXPERT

The prediction system, CRIPES (Chromatographic Retention Index Prediction Expert System) has been implemented using the expert system shell VP-Expert. A brief description of some of the facilities of VP-Expert will be given prior to a discussion of CRIPES.

VP-Expert is an expert system development tool written in Microsoft C to run on an IBM compatible PC with a minimum of 300K of memory. It is a rule based system which operates mainly in the backward chaining mode. There are several features which make it suitable for use with this particular application, the most important of which is probably its capability to handle mathematical routines. The mathematical facilities within VP-Expert include many arithmetic and trigonometric functions. In contrast many other shells are not capable of doing even the simplest calculations unless external high level subroutines are used and appended to the program. VP-Expert can also communicate with compatible external spreadsheets and databases, which meant that the regression coefficients could be held outside the main program and therefore were easily updated. This facility also enables data to be transferred between sections of the program with ease.

The rules are of a standard format:-

```
RULE N
If Y = Z
AND Z > W
OR Z < Z1
THEN X = true;
```

In which W, Y, Z and Z1 are variables or defined values.

The consultation is run from an ACTIONS block which contains the goals which the program must satisfy (e.g. Figure 8.1, Action block for CRIPES). The goals are given using the terminology FIND var. Sub-goals can also be given in the conclusion of the rules or buried within the Rules such that they need to be satisfied before the rule can be evaluated (examples of the form of the Rules in CRIPES are given in Figure 8.2). Variables can be of several types single, plural, or dimensioned. An element of a dimensioned variable may also be defined as a plural variable. The type of variable determines the type of search that VP-Expert undertakes to produce an answer.

During a consultation VP-Expert attempts to satisfy a goal by looking at each rule in turn to find one containing that goal in its conclusion. If the conditions of a rule are satisfied and the variable is a single variable then the goal is satisfied and the consultation moves onto the next goal. If the conditions are not satisfied then the program looks for the next rule containing the goal in its conclusion. The order of the rules in the knowledge base can therefore have a significant influence on the path and results of a consultation because once a goal is satisfied the search for a single variable stops and the consultation moves onto the next goal even if later rules are also satisfied. If after all the rules have been examined and a goal has still not been satisfied then the program checks whether there is an instruction to ask the user for a value (an ASK var:"", Figure 8.1) and then, if such an instruction is found, asks the user to input a value via the keyboard. If no value can be found for the goal, it remains unknown

Figure 8.1: Extract of knowledge base of CRIPES showing ACTIONS block and a selection of ASK and CHOICES statements

```
ACTIONS

PRINTOFF
Display "                    CRIPES

      Chromatographic Retention Index Prediction Expert
System

      This knowledge base will calculate the retention index
      of a single compound from its molecular structure
      at either a single or multiple eluent compositions"
wks rp1,ROW= Benzene,B:Benzmecn
wks rp2,ROW = Benzene,B:Benzmeoh
FIND COLUMN_READ
FIND name
FIND num_aro
FIND check_num
FIND num_ali
FIND subali
FIND SIIaro
FIND SIIali
FIND print
FIND single_eluent
FIND RI;

ASK num_ali:"How many aliphatic substituents present in
{name} ?";
ASK num_aro:"How many aromatic substituents are present in
{name} ?";
ASK subali1:"Which of these aliphatic substituents are
present?";
ASK subali2:"Which of these aliphatic substituents are
present?";
ASK subaro1:"Which, if any, of these aromatic substituents
are present";
ASK subaro2:"Which, if any, of these aromatic substituents
are present";

CHOICES subali2:OH,CONH2,Br,Cl,CN;
CHOICES subali1:CHO,CO2R,OR,COR,CH_CH,ALKYL_CHAIN,ANOTHER;
CHOICES subaro1:COR,CHO,OR,CO2R,CH3,CONH2,OH,CH_CH,ANOTHER;
CHOICES subaro2: NH2,NO2,CN,Cl,Br,Ph;

PLURAL:subali1,subali2,subaro1,subaro2,position,position1;
```

Figure 8.2: Examples of rules used in the knowledge base of CRIPES

```
RULE 1
IF column <> unknown
THEN wks col1,ROW = (column),B:MeCNcol
wks col2,ROW = (column),B:MeOHcol
column_read = done;

RULE 2
IF num_aro <= 6
THEN FIND subs
check_num = ok;

RULE 2d
IF printer = yes
THEN PRINTON
PRINT = done
ELSE
PRINTOFF
PRINT = done;

RULE 3
IF   num_aro > 0
THEN FIND r_a
FIND S1 find s1a
     y=0
     WHILEKNOWN sub2[y]
          y=(y+1)
          find sub3
          FIND S3
          FIND naro
          num = (num + n1)
          FIND check
          FIND rest_alkyl
          FIND pos1
          FIND pos2
          cg1 = (cg1 + (n1 *coeff1[1])+ct1 +CU1)
          cg2 = (cg2 + (n1 *coeff1[2])+ct2 + CU2)
          cg3 = (cg3 + (n1 *coeff1[3])+ct3 + CU3 + R1)
          cg4 = (cg4 + (n1 *coeff2[1])+ct4+ CU4)
          cg5 = (cg5 + (n1 *coeff2[2])+ct5 + CU5)
          cg6 = (cg6 + (n1 *coeff2[3])+ct6 + CU6 + R2)
          RESET coeff1[1] RESET coeff1[2] RESET coeff1[3]
          RESET coeff2[1] RESET coeff2[2] RESET coeff2[3]
          RESET cf[1] RESET cf[2] RESET cf[3] RESET cf2[1]
          RESET cf2[2] RESET cf2[3] RESET ct1 RESET ct2
          RESET ct3 RESET ct4 RESET ct5 RESET ct6
          RESET cU1 RESET cu2 RESET cu3 RESET cu4 RESET cu5
          RESET cu6 RESET c_f RESET w1 RESET position
          RESET position1 RESET pos RESET sub3 RESET pos1
          RESET pos2 RESET pos3
          RESET pos4 RESET pos5 RESET pos6 RESET pos7
          RESET naro Reset n1 RESET s3  RESET w1 RESET c1
          RESET c2 RESET c3 RESET c4 RESET IT
          RESET num_check RESET r_chain RESET rest_alkyl
          RESET r1 RESET r2
 END
          SIIaro = done
ELSE SIIaro=0;
```

and the consultation moves onto the next goal.

If the variable is a plural variable the search is basically the same as for a single variable, except that the program does not stop after the first value of the goal is found but continues to search for as many values as possible.

An ASK statement may be accompanied by a CHOICES var: statement (Figure 8.1). This produces a menu of possible responses from which user selects the required value (or if the variable is a plural variable) values.

The rule base can be written using any text editor such as Wordstar in non-document mode, and VP-Expert also contains a built-in text editor. The program also contains a trace facility which enables the path of a consultation to be monitored and subsequently displayed either as a text file or a graphic decision tree. It is also possible to discover which rule provided a value and why the consultation wants to know a particular value. It is also possible to check the value of any variable at the end of the consultation.

8.5 IMPLEMENTATION AND DEVELOPMENT OF CRIPES

In the previous section a brief description of the facilities and method of working of VP-Expert was given. This shell has been used to implement the retention prediction method, CRIPES, Chromatographic Retention Index Prediction Expert System, using the approach outlined in this section.

The retention index of a compound can be calculated

from the equation

$$RI = PI + \Sigma SI_{Ar-x} + \Sigma SI_{Al-x} + SI_{R} + \Sigma II_{YZ} \qquad 8.1$$

where the terms are as defined earlier (Chapter 3).

The collection of these terms has been described in the previous chapters. Each term in the equation can be described using a quadratic equation. If these are summed they lead to the following equation to describe the retention index

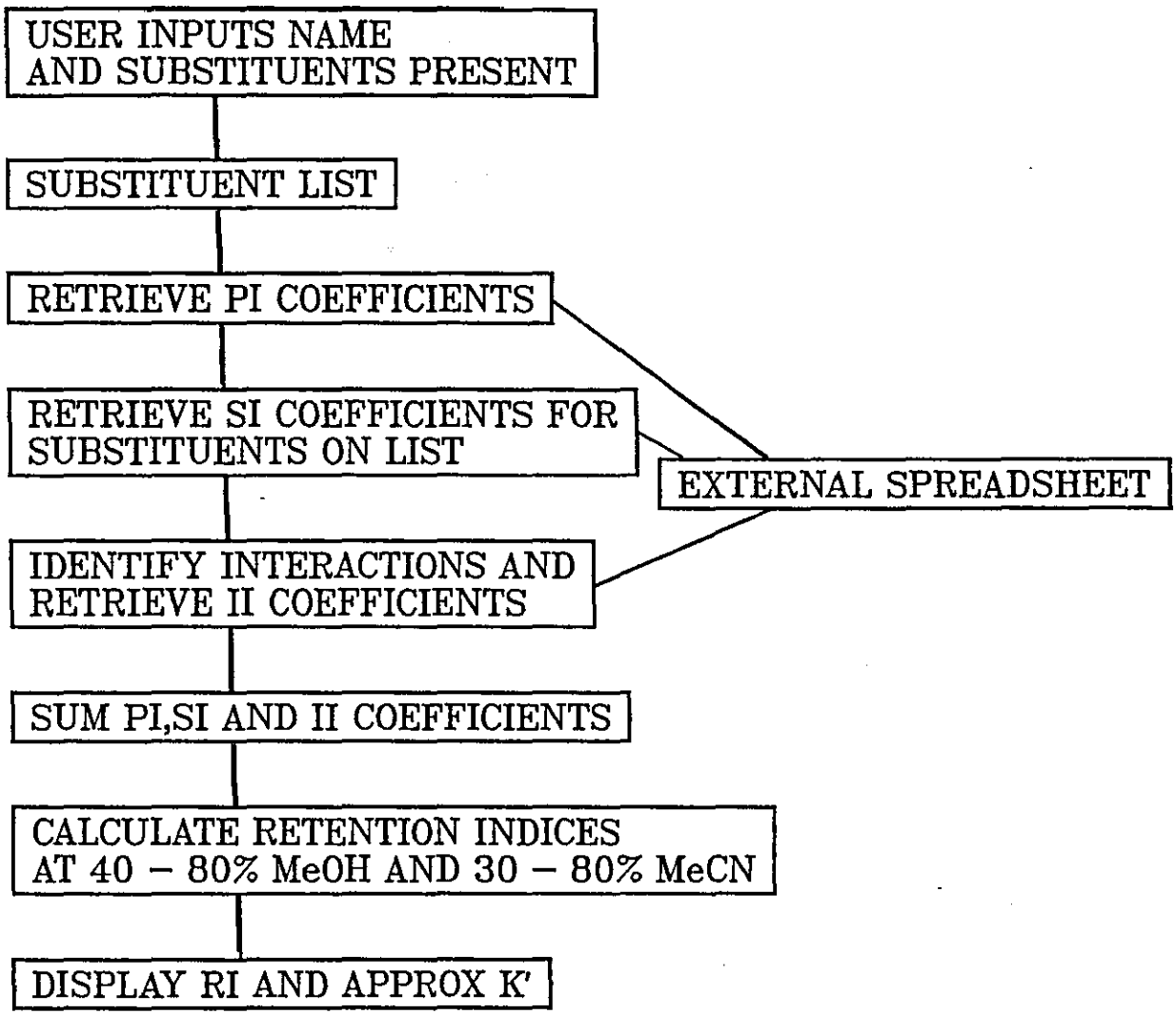
$$RI = x^2\ \Sigma a + x\ \Sigma b + \Sigma c \qquad (x = \%\ \text{modifier}) \qquad 8.2$$

enabling the calculation of retention index at any eluent composition.

The approach used to obtain the coefficients of Equation 8.3 is shown in the flowchart in Figure 8.3. Two equations are obtained for each compound, one describing the change with methanol proportion and the other with acetonitrile proportion. All the regression coefficients of the regression equations for *PI*, *SI*, and *II* were held in the external spreadsheets and accessed through the expert system. A summary of the values of the coefficients of the equations used within CRIPES and held in the spreadsheets is given in Tables 8.1 to Table 8.6.

The aromatic and aliphatic substituents were treated separately. The user first enters the aromatic substituents selected from menus presented by CRIPES, then the aliphatic substituents. Then taking each aromatic substituent in turn the program asked the user to input the number of that substituent present in the compound and in turn the

Figure 8.3: Flowchart showing mode of operation of CRIPES



position/positions relative to a hydroxyl, amino or alkyl group if these are present. This enabled CRIPES to extract the coefficients for the substituent indices and to identify any interactions and extract the appropriate coefficients for these. The coefficients are summed to provide an overall equation describing the aromatic contribution. Having emptied the list of aromatic substituents the program then repeats the process for the aliphatic substituents. The program prompts the user for information on the branching of an alkyl chain, positions of aliphatic substituents relative to the aromatic ring and for the length of the alkyl chain in mixed alkyl-aryl groups such as PhCOR and $PhCO_2R$.

The coefficients for the aromatic contribution, the aliphatic contribution and the parent contribution are then summed to give two equations relating the retention index of the compound to eluent composition. One equation describing changes in methanol and the other in acetonitrile. Using the equations the program then calculates the retention index values over the ranges 40 - 80% methanol and 30 - 80% acetonitrile at 10% intervals. Using the program it is therefore possible to calculate the retention index of any substituted benzene directly from its molecular structure.

The final stage of the program is to calculate an approximate capacity factor. This enables an estimated retention time to be obtained to give an indication of the length of an analysis. This calculation is based on the relationship:-

$$\log k' = a'RI + b' \qquad 8.3$$

The coefficients a' and b' in Equation 8.3 are known from

Table 8.1: Coefficients of regression equations for parent index equations held in spreadsheet and accessed by CRIPES

$$PI = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Parent	Organic modifier	Coefficients of parent index equations a	b	c
benzene	methanol	-0.0121	3.887	748
benzene	acetonitrile	-0.0154	2.761	841

Table 8.2: Coefficients of regression equations for substituent index equations for aliphatic substituents (SI_{Al-x}) held in the spreadsheet accessed by CRIPES

$$SI = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Substituent	Coefficients of substituent index equations Methanol a	b	c	Acetonitrile a	b	c
$CONH_2$	0.0079	-5.013	-178	0.0855	-11.786	-179
prim-OH	-0.0257	-0.494	-273	0.0561	-8.139	-223
sec-OH	-0.0257	-0.894	-286	0.0346	-5.979	-310
CN	-0.0250	-1.050	-185	0.0002	-2.997	-160
CHO	0.1314	-18.531	337	0.0073	-1.768	-184
CO_2R	0.0071	-3.717	-90	0.0130	-3.580	-164
COR	-0.0086	-1.851	-201	0.0161	-3.654	-206
OR	0.0136	-2.939	-154	0.0211	-2.644	-218
Cl	-0.0021	-1.053	19	0.0018	-1.962	27
Br	-0.0536	4.719	-99	0.0325	-4.792	121
CH:CH	-0.0314	2.891	55	0.0100	-1.780	154

Table 8.3: Coefficients of regression equations for substituent indices for aromatic substituents (SI_{Ar-x}) held spreadsheet and accessed by CRIPES

$$SI = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Substituent	Coefficients of substituent index equations					
	Methanol			Acetonitrile		
	a	b	c	a	b	c
CH_3	0	0	100	0	0	100
$CONH_2$	0.0093	-4.804	-104	0.1260	-14.878	2
NH_2	-0.0264	0.541	-215	0.0118	-2.405	-153
OH	-0.0271	0.117	-167	0.0218	-4.616	-93
CHO	0.0186	-4.469	39	0.0025	-1.335	-92
CN	-0.0114	-1.429	-34	-0.0025	-1.251	-57
COR	0.0114	-3.791	-52	0.015	-2.704	-143
NO_2	0.0050	-2.390	53	-0.0104	-0.586	-14
OR	0.0129	-2.263	-30	0.0029	-1.097	-80
CO_2R	0.0143	-3.774	43	0.0105	-2.096	-67
Cl	0.0086	-1.669	167	0	0	98
Br	0.015	-2.190	207	0	0	127
Ph	0.025	-3.870	436	0.0193	-3.299	372
CH:CH	-0.0307	2.956	74	0.0223	-2.741	189

Table 8.4: Coefficients of interaction index equations held in spreadsheet and accessed by CRIPES - aliphatic interaction index equations

In the table I- represent substituents on a benzylic carbon ($PhCH_2$-X) and II- substituents on $PhCH_2CH_2$-X

$$II = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Substituent	Coefficients of interaction index equations					
	Methanol			Acetonitrile		
	a	b	c	a	b	c
I-$CONH_2$	0.0186	-1.809	56	0.0018	0.395	6
I-OH	-0.0007	0.996	-3	-0.0046	1.236	29
I-CN	0.0221	-2.327	113	0	0.440	52
I-CO_2CH_3	0	0	0	0	0	33
I-$COCH_3$	0	0	-12	0	0	10
I-Cl	-0.0143	1.414	-31	-0.0079	0.479	5
I-Br	0.0836	-9.139	235	0	0	17
II-CO_2CH_3	0	0	17	0	0	24
II-OH	-0.0079	1.233	-16	-0.0021	0.636	10
II-Cl	0	0	-12	-0.0075	0.482	-10
II-Br	-0.0071	0.757	-26	-0.0101	0.668	-14
II-CN				0	0	19
branch	0	0	-12	0	0	-20

Table 8.5: Coefficients of interaction index equations held in spreadsheet and accessed by CRIPES - aromatic interaction index equations for interactions in methanol eluents

In the table T represent the interaction with alkyl or methyl groups and P the interactions with phenolic hydroxyl groups.
2, 3, and 4 are the position of the substituent (relative to the methyl or phenol group) and if no number is listed the interaction index applies to all positions. Where the 3- substituent is marked * the interaction index applies to both 3- and 4- positions.

$II = ax^2 + bx + c$ x = % modifier

Substituent	Coefficients of *II* equations			Substituent	Coefficients of *II* equations		
	a	b	c		a	b	c
T-Br	0.0150	-1.230	32	T2-NO_2	0	0	-28
T-Cl	0.0150	-1.230	32	T3-NO_2	0.0036	-0.119	-2
T2-CH_3	0.0221	-1.967	31	T4-NO_2	0	0	-12
*T3-CH_3	0.0207	-1.796	45	T2-Ph	0	0	-27
T2-OCH_3	0.0057	-0.306	23	T4-Ph	0	0	21
T2-CO_2CH_3	0	0	-11	T2-$CONH_2$	-0.0429	5.203	-208
T-NH_2	0	0	-13	*T3-$CONH_2$	0	0	-13
P2-NH_2	0.0471	-2.577	161	P2-$CONH_2$	-0.0214	3.551	147
P3-NH_2	0.0543	-3.274	78	P4-$CONH_2$	0	1.550	8
P4-NH_2	0.1707	-15.756	363	P2-CN	-0.1230	10.445	-161
P2-Cl	-0.0343	3.574	-96	P3-CN	0.0150	-0.730	125
*P3-Cl	-0.0171	1.777	-7	P4-CN	-0.0725	7.205	-97
P2-Br	-0.0550	5.790	-155	P2-OCH_3	0.0079	-0.313	36
*P3-Br	-0.0171	1.777	-7	P3-OCH_3	0	0	18
P2-OH	0.0736	-5.439	216	P4-OCH_3	0	0	-16
P3-OH	0.0621	-4.787	142	*P3-CH_3	0.0114	-1.811	51
P4-OH	0.0779	-5.493	87	P2-NO_2	0	-2.000	200
P2-CHO	0.0129	1.017	154	P3-NO_2	0	0.210	96
P3-CHO	-0.0021	1.067	49	P4-NO_2	0	-3.800	145
P4-CHO	0	-2.150	122	P2-Ph	0	0	-9
P2-CO	0.0321	0.253	193	*P3-Ph	0.0076	-1.313	46
P3-$COCH_3$	0.0420	-3.627	159	P3-CO_2CH_3	-0.0179	2.493	-11
P4-$COCH_3$	-0.1240	12.914	-274	P4-CO_2CH_3	-0.1057	11.186	-231

Table 8.6: Coefficients of interaction index equations held in spreadsheet and accessed by CRIPES - aromatic interaction index equations for interactions in acetonitrile eluents

In the table T represent the interaction with alkyl or methyl groups and P the interactions with phenolic hydroxyl groups.
2,3 and 4 are the position of the substituent (relative to the methyl or phenol group) and if no number is listed the interaction index applies to all positions.

$$II = ax^2 + bx + c \qquad x = \%\ \text{modifier}$$

Substituent	Coefficients of *II* equations a	b	c	Substituent	Coefficients of *II* equations a	b	c
T-Br	0.0232	-2.182	50	T2-CHO	0.0073	-0.454	-7
T-Cl	0.0180	-1.875	48	T3-CHO	-0.0025	0.400	-14
T2-CH_3	0	0	-22	T4-CHO	0.0075	-0.428	-16
T2-$COCH_3$	0	0	-11	T2-$CONH_2$	0	0	-58
*T3-$COCH_3$	0	0	-19	T3-$CONH_2$	-0.0288	2.631	-82
T2-OCH_3	0.0062	-0.325	18	T4-$CONH_2$	0	0	-29
T3-OCH_3	0	0	-15	T-NH_2	0	0	-27
T2-CO_2CH_3	0	0	-16	T2-CN	0	0	-20
*T3-CO_2CH_3	0	0	-10	T4-CN	0	0	-15
T2-NO_2	0	0	-25	T2-Ph	0	0	-21
T4-NO_2	0	0	-12				
P2-NH_2	0.0511	-1.124	94	P2-CN	-0.0757	7.540	-117
P3-NH_2	0.0388	-2.525	88	P3-CN	0.0092	-0.604	122
P4-NH_2	0.0759	-5.794	103	P4-CN	0.0909	-7.044	10
P2-Br	-0.0134	0.670	-12	P2-OCH_3	-0.0027	1.366	-6
*P3-Br	-0.0052	-0.490	53	P3-OCH_3	-0.0291	2.973	-48
P2-Cl	-0.0018	-0.798	66	P4-OCH_3	-0.0309	3.373	-95
*P3-Cl	-0.0052	-0.490	53	P2-NO_2	-0.0486	7.877	-88
P2-OH	-0.0386	6.737	-72	P3-NO_2	-0.0350	4.356	-16
P3-OH	0.0145	0.620	20	P4-NO_2	-0.1314	12.646	-278
P4-OH	0.0246	0.455	-36	P2-CH_3	-0.0105	0.816	-25
P2-CHO	0.0454	-3.000	355	*P3-CH_3	-0.0030	-0.109	-18
P3-CHO	-0.0100	1.380	48	P2-Ph	-0.0477	4.799	-104
P4-CHO	-0.0313	3.089	-37	*P3-Ph	-0.0198	1.266	-26
P2-CO	-0.0311	5.598	121	P2-$CONH_2$	-0.1500	16.937	-114
P3-$COCH_3$	0.0059	0.128	65	P4-$CONH_2$	0	1.910	8
P4-$COCH_3$	0.0071	0.171	20				
P3-CO_2CH_3	0.0177	-2.033	108				
P4-CO_2CH_3	-0.0157	1.783	-9				

the experimental regression equations for the alkylarylketone standards relating carbon number to log k'. The values of the coefficients a' and b' were dependent on the eluent composition and the change with organic modifier concentration was described by quadratic regression equations (Chapter 3, Table 3.23). Using the quadratic regression equations relating a' and b' to eluent composition it was possible to calculate k' at any eluent composition. One advantage of this facility is that as retention indices should be largely independent of the brand of column it should be possible to predict the capacity factors on any ODS-silica column if the retentions of the alkylarylketones on that column are known. So far CRIPES accesses a spreadsheet which contains details of a' and b' for both Spherisorb-ODS2, from this study, and Hypersil-ODS calculated from earlier work in this laboratory[134,145] (in acetonitrile, intercept a = 2.36×10^{-4}, b = 0.0390 and c = 3.027, and slope a = 0.36×10^{-6}, b = -8.08×10^{-5} and c = 5.4×10^{-3}, in methanol, intercept a = 1.98×10^{-4}, b = -0.0106 and c = 2.188, and slope a = 3.3×10^{-6}, b = 4.5×10^{-5} and c = 5.58×10^{-3}).

8.6 CALCULATION OF RESOLUTION

Using the regression equations for calculating retention index, in different eluents, it should be possible to calculate the optimum separation between two compounds and therefore suggest suitable separation conditions. Although it is very easy to calculate the maximum difference in retention indices between two compounds, this value does not correspond with the maximum resolution calculated by

Equation 8.4.

$$R_m = 2\ (t_{R2} - t_{R1})/(w_1 + w_2) \qquad 8.4$$

Where R_m is the resolution, t_{R1} and t_{R2} the retention times of the first and second compounds and w_1 and w_2 the width at the base of the first and second peaks. The resolution varies linearly with the retention and the retention index logarithmically. It was therefore necessary to calculate the separation in terms of the capacity factors.

Using CRIPES the retention index is related to eluent composition using the equation

$$RI = Ax^2 + Bx + C \qquad 8.5$$

Where $A = \Sigma a$, $B = \Sigma b$ and $C = \Sigma c$.

The retention index is related to the capacity factor using Equation 8.3 at a particular eluent composition.

$$\log k' = a'RI + b' \qquad 8.3$$

The coefficients a' and b' are also related to eluent composition using the Equations 8.6 and 8.7:-

$$a' = a_1x^2 + b_1x + c_1 \qquad 8.6$$

$$b' = a_2x^2 + b_2x + c_2 \qquad 8.7$$

Combining the equations 8.3, 8.5, 8.6 and 8.7 the overall equation describing the change in k' with eluent composition

is:-

$$\log k' = (a_1x^2 + b_1x + c_1)(Ax^2 + Bx + C) + a_2x^2 + b_2x + c_2 \qquad 8.8$$

This gives Equation 8.9:-

$$\log k' = a_1Ax^4 + (a_1B + Ab_1)x^3 + (a_1C + b_1B + c_1A + a_2)x^2 + (b_1C + c_1B + b_2)x + (c_1C + c_2) \qquad 8.9$$

Collecting the constants to simplify the equation we have:-

$$\log k' = A_1x^4 + B_1x^3 + C_1x^2 + D_1x + E_1 \qquad 8.10$$

Previous work on capacity factors has suggested that the relationship between log k' and eluent composition was quadratic[29], in this case we have derived a more complex polynomial. This suggests that even the quadratic equation may be a simplification of the true situation although in practice the x^4 and x^3 terms may prove insignificant.

By using the equation above it should be possible to calculate the capacity factor at any eluent composition, it should therefore be possible to calculate the resolution between two compounds by comparing the equations to calculate capacity factors.

The resolution (R_m) of two closely eluting compounds (1 and 2) can be described using the equation

$$R_m = 1/4\ (\alpha/\alpha-1)(k'_2/\ 1 + k'_2)\ \sqrt{N} \qquad 8.11$$

where $\alpha = k_2/k_1$ and N the number of theoretical plates. Having obtained this equation the next stage was to

incorporate it into CRIPES.

CRIPES contains equations for log k', however resolution varies [illegible] with k' not log k' [illegible]. It was therefore necessary to take the antilog of the equation to obtain k', although this should be straightforward the equation for log k' was in $\log_{10}$ and VP-Expert can only work with natural logs, k' for the first compound was therefore calculated using the equation

$$k'_1 = \exp((A_1x^4 + B_1x^3 + C_1x^2 + D_1x + E) \times 2.30259) \qquad 8.12$$

and an equivalent equation for compound 2. Substitution of these equations into the equation for resolution gives the equation for resolution at any eluent composition.

The aim in optimising the resolution should be to obtain the maximum resolution between the two compounds. This could be done by differentiating the equation and solving it to find the turning point where the resolution is a maximum. However, this presented problems as the solution is beyond the mathematical capabilities of VP-Expert. In addition, in practice the separation required may not be the maximum resolution but a separation which can be obtained within defined limits of eluent compositions and capacity factors.

It was therefore necessary to take a simpler more direct approach, consequently CRIPES calculates the retention index equations for the two compounds, in turn, and stores them in a database file. These equations are then used to calculate the capacity factors and resolution at 10% intervals over the eluent ranges and the program then displays the change in resolution with eluent composition.

The user can select a range to be expanded if required, the resolution would then be calculated at 1% intervals. The selection of optimum experimental conditions would then be made manually taking into account the length of analysis and the resolution required.

If it is found desirable to calculate the maximum resolution a possible approach would be to use a second equation solving program such as Eureka: the solver, Borland International, which could find the true optimum resolution. It would be possible to link this to the expert system by using spreadsheets to transfer the data between the programs.

8.7 TESTING OF CRIPES

The program CRIPES has been tested in two ways. Firstly the retention indices of a number of model compounds have been calculated to check that the program was capable of extracting the appropriate data from the spreadsheets. These results matched in each case. Secondly, the retention of a number of, usually polysubstituted, test compounds containing the groups which have been studied previously, have been measured at selected eluent compositions and compared to those calculated by CRIPES both as retention indices (Tables 8.7 and 8.8) and capacity factors (Tables 8.9 and 8.10). These compounds were not included in the data used to calculate the substituent indices and interaction indices.

The compounds included in the trials include several which were selected to test specific aspects of the

retention prediction. Firstly, there were a number of compound including the grouping $-COCH_2Br$, including phenacylbromide. These were used to determine whether an aliphatic substituent on an aliphatic chain attached to an aromatic substituent had the same substituent index as when it was substituted on an alkylbenzene. Secondly a limited number of aromatic compounds were included in which there were amino groups plus another substituent. The retention indices of these compounds were calculated using the hydroxyl interaction terms to determine whether the interaction indices were closely related. The remaining compounds covered a range of possible interactions and each of these groups will be discussed in turn.

In all the compounds an assumption of additivity of the interaction indices has been made. All the possible interaction indices are summed and there is no judgement exercised as to whether one, or more, of the interactions would dominate to the exclusion of other interactions. Also the interaction terms used do not include any interactions between aliphatic substituents or interactions between substituents on an alkyl side chain and the aromatic substituents as so far these parameters have not yet been determined.

Generally the discussion will concentrate on the correlation between retention indices and the capacity factors given as examples. With the capacity factors when the value is small the errors in between the calculated and experimental values will also be small in absolute terms.

Table 8.7: Experimental retention indices (RI_m) and retention indices calculated (RI_c) by CRIPES for a selection of test compounds

Compound	Retention index Methanol proportion (%) 40		50		60		70		80	
	RI_m	RI_c	RI_m	RI_c	RI_m	RI_c	RI_m	RI_c	RI_m	RI_c
phenacyl bromide			860	802	859	790	847	768	807	735
α-bromo-p-phenylacetophenone					1169	1081			1120	1056
α-p-dibromoacetophenone			1009	945	1005	932	993	915	971	892
4-nitrophenacyl bromide			858	748	855	718	834	678	782	629
α-chloro-3,4-dihydroxy acetophenone			547	513	547	480	414	458	161	448
o-bromoaniline					840	914	840	890	818	853
m-bromoaniline					814	828	804	815	752	794
o-nitroaniline					764	668	760	625	711	575
m-nitroaniline					695	696	687	675	649	647
p-nitroaniline					648	505	642	444	556	376
benzylacetate	891	857	889	850	891	845	881	845	875	840
benzyl 2-bromoacetate			956	857	948	841	934		919	775
benzyl chloromethyl ether			935	806	901	805	909	803	969	801
1-bromo-2-nitrobenzene			931	993	930	995	921	998	889	1003
2-bromo-4-methylphenol					895	922	874	916	830	908
t-butylhydroquinone					771	799			637	786
p-t-butylphenol					998	1003	970	980	933	950
2-chloromethyl-4-nitrophenol			427	695	507	662				
4,6-dichloro-1,3-dihydroxybenzene					695	687			(307)	594
3,4-dimethoxyacetophenone					729	761			695	742
N,N-dimethylbenzamide	717	803	709	791	694	779	683	766	665	753
2,6-dimethyl-4-nitrophenol			746	830	708	812	559	787		
2,4-dimethylphenol			863	872	838	864	848	849	792	829
2,5-dimethylphenol			860	872	863	864	844	849	822	830
dimethyl phthalate					785	874	757	859	716	847
N-ethylaniline			875	858	886	860	894	854	879	841
ethyl benzoate			995	1002	995	1006	998	1010	1004	1014
ethyl 3-phenylpropionate			1046	1059	1045	1055	1040	1050	1037	1045
ethylphenylacetate			947	954	947	950	930	945	907	940
ethyl phenylcyanoacetate	860	745	733	706	777	665	708	623	608	579
2-hydroxybenzyl alcohol					605	439			554	346
4-hydroxybenzyl alcohol					519	439			511	346
2-hydroxy-5-nitrobenzyl bromide			429	720	510	690				
N-methylbenzamide			636	691	623	679	616	666	591	653
m-nitrobenzylalcohol					688	628	688	608	607	582
p-nitrobenzylalcohol					678	612	665	588	601	559
4-phenyl-1-butanol					912	931	888	915	841	893
5-phenyl-1-pentanol					1003	1031	979	1015	937	993
n-propyl p-hydroxybenzoate	940	893	914	873	882	848	844	818	809	783
thymol	1042	1043	1035	1038	1030	1027	1001	1009	965	984

Table 8.7: Experimental retention indices (RI_{exp}) and retention indices calculated (RI_{calc}) by CRIPES for a selection of test compounds

Compound	Retention index Acetonitrile proportion (%) 40		50		60		70		80	
	RI_e	RI_c	RI_e	RI_c	RI_e	I_c	RI_e	RI_c	RI_e	RI_c
phenacyl bromide			891	763	878	751	857	745	832	746
α-bromo-p-phenylacetophenone					1135	994			1075	977
α-p-dibromoacetophenone			1028	888	1015	880	994	883	972	896
4-nitrophenacyl bromide			898	693	874	664	830	639	784	618
α-chloro-3,4-dihydroxy acetophenone			593	449			538	393		
o-bromoaniline			866	939	860	930	855	919	846	903
m-bromoaniline			841	839	828	828	820	816	794	801
o-nitroaniline			778	812	762	820	748	815	720	797
m-nitroaniline			742	742	723	729	709	707	668	675
p-nitroaniline			701	653	669	618	663	553	639	459
benzylacetate	883	873	877	863	869	852	859	841	844	829
benzyl 2-bromoacetate	990	855					922	786	886	775
benzyl chloromethyl ether	645	788					624	758	629	750
2-bromo-4-methylphenol			863	897	846	889	823	887	812	892
1-bromo-2-nitrobenzene			947	998	934	992	919	980	899	963
t-butylhydroquinone					978	768			974	783
p-t-butylphenol			943	974	924	958	905	943	882	929
2-chloromethyl-4-nitrophenol			432	642	567	592	332	540	(281)	484
4-chloro-2-nitroaniline					881	925	868	906	846	876
4,6-dichloro-1,3-dihydroxybenzene					627	697	601	673	574	658
3,4-dimethoxyacetophenone					696	729			697	702
N,N-dimethylbenzamide	653	735	631	714	633	714	643	737	683	781
2,4-dimethylphenol			831	830	777	811	803	799	744	768
2,5-dimethylphenol			829	830	817	811	795	799	783	768
dimethyl phthalate			797	850	784	841	764	834	750	828
2,6-dimethyl-4-nitrophenol			770	791	743	759	699	724	528	686
N-ethylaniline			929	897	928	896	940	895	925	893
ethyl benzoate			977	995	987	996	991	997	994	996
ethyl 3-phenylpropionate	1036	1052	1030	1042	1021	1031	1011	1020	999	1008
ethylphenylacetate			947	963	940	952	921	941	905	929
ethylphenylcyanoacetate	914	763					819	655	754	618
2-hydroxybenzylalcohol					796	350			801	297
4-hydroxybenzylalcohol					520	337			463	324
2-hydroxy-5-nitrobenzyl bromide			353	663	303	615	264	569	(233)	525
N-methylbenzamide	589	635	560	614	552	614	549	637	558	682
p-nitrobenzyl alcohol			638	549	628	530	624	516	599	507
m-nitrobenzyl alcohol			668	561	643	542	541	528	611	519
4-phenyl-1-butanol			821	856	801	830	807	829	797	836
5-phenyl-1-pentanol			903	956	883	930	886	929	879	936
n-propyl p-hydroxybenzoate			862	850			779	787	775	773
thymol	1001	1008	984	991	968	972	955	951	941	929

Table 8.8: Experimental capacity factors (k'_m) and capaci ty factors estimated (k'_c) by CRIPES

Compound	Capacity factor Methanol proportion (%) 40		50		60		70		80	
	k'_m	k'_c	k'_m	k'_c	k'_m	k'_c	k'_m	k'_c	k'_m	k'_c
phenacyl bromide			3.94	2.99	2.48	1.52	1.13	0.82	0.69	0.48
α-bromo-p-phenylacetophenone					15.92	9.00			1.96	1.30
α-p-dibromoacetophenone			10.91	8.55	5.95	3.59	2.19	1.64	1.20	0.83
4-nitrophenacyl bromide			3.88	2.01	2.42	0.98	1.07	0.53	0.64	0.33
α-chloro-3,4-dihydroxy acetophenone			0.47	0.36	0.38	0.23	0.16	0.19	0.08	0.18
o-bromoaniline					2.21	3.21	1.10	1.45	0.72	0.72
m-bromoaniline					1.89	1.91	0.94	1.02	0.57	0.59
o-nitroaniline					1.40	0.72	0.77	0.41	0.50	0.28
m-nitroaniline					0.93	0.86	0.55	0.52	0.41	0.35
p-nitroaniline					0.70	0.27	0.45	0.17	0.30	0.14
benzylacetate	13.88	9.52	5.51	4.38	2.76	2.18	1.41	1.18	0.76	0.69
benzyl 2-bromoacetate			8.71	4.48	3.64	2.07	1.81	1.02	0.89	0.55
benzyl chloromethyl ether			7.47	3.09	2.98	1.66	1.61	0.96	1.07	0.60
1-bromo-2-nitrobenzene			6.38	12.18	3.79	5.24	1.58	2.44	0.91	1.22
2-bromo-4-methylphenol					3.08	3.37	1.28	1.65	0.75	0.88
t-butylhydroquinone					1.46	1.61			0.39	0.57
p-t-butylphenol					5.72	5.49	1.96	2.24	1.05	1.02
2-chloromethyl-4-nitrophenol			0.21	1.37	0.30	0.70				
4,6-dichloro-1,3-dihydroxybenzene					0.93	0.81			(0.13)	0.29
3,4-dimethoxyacetophenone					1.13	1.27			0.48	0.49
N,N-dimethylbenzamide	2.82	5.97	1.46	2.76	0.85	1.42	0.55	0.81	0.36	0.51
2,6-dimethyl-4-nitrophenol			1.82	3.68	1.00	1.73	0.30	0.89		
2,4-dimethylphenol			4.03	5.01	2.18	2.37	1.14	1.20	0.66	0.67
2,5-dimethylphenol			3.94	5.01	0.73	2.37	1.12	1.20	0.73	0.67
dimethyl phthalate					1.58	2.52	0.76	1.25	0.51	0.71
N-ethylaniline			4.36	4.52	2.91	2.32	1.40	1.23	0.96	0.69
ethyl benzoate			11.76	13.01	4.80	5.60	2.44	2.58	1.27	1.27
ethyl 3-phenylpropionate			16.84	19.75	6.44	7.55	2.99	3.13	1.38	1.41
ethylphenylacetate			7.11	9.14	4.21	4.00	1.65	1.90	0.97	0.98
ethyl phenylcyanoacetate	9.92	3.62	1.72	1.48	1.32	0.71	0.62	0.41	0.29	0.28
2-hydroxybenzyl alcohol					0.54	0.18			0.30	0.12
4-hydroxybenzyl alcohol					0.32	0.18			0.26	0.12
2-hydroxy-5-nitrobenzyl bromide			0.21	1.64	0.30	0.83				
N-methylbenzamide	1.46	2.52	0.86	1.33	0.55	0.78	0.40	0.50	0.28	0.36
m-nitrobenzyl alcohol					0.89	0.57	0.56	0.38	0.36	0.28
p-nitrobenzyl alcohol					0.83	0.52	0.50	0.35	0.35	0.26
4-phenyl-1-butanol					3.40	3.55	1.36	1.64	0.78	0.83
5-phenyl-1-pentanol					5.87	6.50	2.03	2.65	0.36	1.18
n-propyl p-hydroxybenzoate	20.02	13.06	6.56	5.05	2.46	2.16	1.18	1.03	0.62	0.57
thymol	62.29	47.53	12.99	16.97	6.91	6.37	2.26	2.56	1.17	1.14

Table 8.9: Experimental capacity factors (k'_m) and capacity factors estimated (k'_c) by CRIPES

Compound	Capacity factor									
	Acetonitrile Concentration (%)									
	40		50		60		70		80	
	k'_m	k'_c	k'_m	k'_c	k'_m	k'_c	k'_m	k'_c	k'_m	k'_c
phenacyl bromide			3.24	1.51	1.69	0.94	0.90	0.64	0.67	0.47
α-bromo-p-phenylacetophenone					4.92	2.53			1.33	0.92
α-p-dibromoacetophenone			4.95	2.87	2.99	1.59	1.40	1.02	0.99	0.73
4-nitrophenacyl bromide			2.61	1.06	1.66	0.66	0.82	0.45	0.58	0.33
α-chloro-3,4-dihydroxy acetophenone			0.61	0.31			0.32	0.20		
o-bromoaniline			2.41	3.71	1.57	1.95	0.97	1.14	0.70	0.74
m-bromoaniline			2.14	2.24	1.37	1.29	0.87	0.81	0.60	0.55
o-nitroaniline			1.57	1.95	1.04	1.24	0.69	0.81	0.49	0.55
m-nitroaniline			1.33	1.36	0.89	0.86	0.61	0.56	0.42	0.38
p-nitroaniline			0.52	0.87	0.71	0.55	1.09	0.34	0.39	0.21
benzylacetate	5.95	5.00	2.87	2.53	1.64	1.42	0.99	0.88	0.57	0.60
benzyl 2-bromoacetate	9.91	4.43					1.23	0.73	0.64	0.51
benzyl chloromethyl ether	1.07	2.89					0.43	0.67	0.30	0.48
2-bromo-4-methylphenol			2.38	3.00	1.48	1.65	0.88	1.03	0.63	0.72
1-bromo-2-nitrobenzene			3.30	5.03	2.13	2.51	1.10	1.40	0.81	0.88
t-butylhydroquinone					2.56	1.01			1.00	0.52
p-t-butylphenol			3.51	4.45	2.04	2.19	1.15	1.24	0.77	0.80
2-chloromethyl-4-nitrophenol			0.28	0.82	0.46	0.49	0.16	0.32	(0.14)	0.22
4-chloro-2-nitroaniline					2.04	1.89	1.02	1.10	0.70	0.68
4,6-dichloro-1,3-dihydroxybenzene					0.59	0.76	0.39	0.51	0.32	0.36
3,4-dimethoxyacetophenone					0.79	0.86			0.46	0.41
N,N-dimethylbenzamide	1.12	2.06	0.81	1.18	0.60	0.81	0.46	0.62	0.35	0.52
2,4-dimethylphenol			1.90	2.13	1.11	1.20	0.75	0.74	0.52	0.50
2,5-dimethylphenol			2.02	2.13	1.31	1.20	0.80	0.74	0.58	0.50
dimethyl phthalate			1.73	2.36	1.14	1.36	0.73	0.86	0.53	0.60
2,6-dimethyl-4-nitrophenol			1.42	0.82	0.96	0.49	0.54	0.32	0.28	0.22
N-ethylaniline			3.03	3.00	2.08	1.70	1.18	1.06	0.87	0.72
ethylbenzoate			5.07	4.95	2.70	2.55	1.57	1.48	0.83	0.97
ethyl 3-phenylpropionate	13.39	15.76	6.18	6.29	3.12	2.94	1.69	1.60	0.89	1.00
ethylphenylacetate			3.30	4.20	2.18	2.13	1.11	1.23	0.82	0.80
ethyl phenylcyanoacetate	6.09	2.47					0.86	0.48	0.43	0.32
2-hydroxybenzyl alcohol					1.20	0.17			0.61	0.13
4-hydroxybenzyl alcohol					0.38	0.17			0.24	0.13
2-hydroxy-5-nitrobenzyl bromide			0.31	0.91	0.27	0.54	0.16	0.36	(0.12)	0.25
N-methylbenzamide	0.74	1.08	0.56	0.71	0.43	0.54	0.33	0.45	0.24	0.39
p-nitrobenzyl alcohol			0.80	0.51	0.60	0.38	0.46	0.30	0.35	0.24
m-nitrobenzyl alcohol			0.92	0.54	0.63	0.40	0.50	0.31	0.36	0.24
4-phenyl-1-butanol			1.95	2.23	1.22	1.30	0.83	0.85	0.61	0.61
5-phenyl-1-pentanol			2.38	3.72	1.72	1.95	1.08	1.18	0.76	0.82
n-propyl p-hydroxybenzoate	4.33	4.30					0.74	0.74	0.44	0.51
thymol	12.73	11.91	4.47	4.84	2.46	2.31	1.35	2.31	0.91	0.80

a) Phenacyl halides

The calculated retention indices ($RI_{=}$) of phenacyl bromides and chlorides were significantly smaller than the experimental values (RI_{∞}). These phenacyl halogens are chemically different to the corresponding alkyl halides and generally much more reactive. These observations suggest that a constant interaction was occurring with the carbonyl and that probably an interaction index term could be included into the expert system.

Benzyl 2-bromoacetate may also be included in this section, in this compound the experimental retention indices were larger than the calculated values. The interaction was of a similar magnitude to that observed with the phenacyl halides although a separate term might be required.

b) Compounds with amino groups

Five substituted anilines were studied in which the hydroxyl interaction indices were assumed to apply to amino groups in the related position on the aromatic ring to determine whether the interactions were of the same magnitude. These compounds were *ortho*- and *meta*-bromoaniline and *ortho, meta, and para*-nitroaniline. In both methanol and acetonitrile containing eluents the retention indices of *m*-bromoaniline and *m*-nitroaniline were close to the calculated values. However for the *ortho* and *para* isomers there was a poor correlation. The deviations of the *ortho* substituents from the calculated values suggests that the interactions with the amino group may be very different in magnitude to

those of the hydroxyl group when the substituents are in close proximity. In Chapter 7 the possibility of ionisation of the model compound used to calculate the interaction index for *para* hydroxy-nitro. The difference between the experimental and calculated *RI* for *p*-nitroaniline values may support this and suggests that the interaction index may not be reliable.

c) Amide test compounds

The secondary and tertiary amides, N-methylbenzamide and N,N-dimethylbenzamide showed relatively large deviations from the retention indices calculated on the basis of the primary amide, (about 80 units for N,N-dimethylbenzamide and 40 to 130 for N-methylbenzamide). A separate substituent index would be required to account for these substituents and substituent indices for the functional groups CONHR and CONHR,R' could be added to CRIPES.

d) Secondary amine

Unlike the secondary amide the calculated and experimental retention indices of the secondary amine, N-ethylaniline, were fairly close (about 30 units difference). The secondary amine apparently behaves like the primary amine and a separate substituent index would probably not be required, although a small interaction term might be necessary.

e) Aliphatic substituted compounds

The two alcohols 4-phenyl-1-butanol and 5-phenyl-1-propanol were included to test the validity of the assumption that the interaction with the ring was not significant at a carbon chain length greater than 2. In both methanol and acetonitrile there were differences between the experimental and calculated retention indices of about 30 units. For these compounds the interactions with the ring may be significant over longer distances although further work would have to be done to confirm the significance of the interactions.

The aliphatic disubstituted compounds discussed earlier (Chapter 5) showed that interactions can occur between aliphatic substituents and that it will be necessary in future work to consider a range of disubstituted compounds to derive interaction index terms. The disubstituted compounds included in this data set (benzyl chloromethyl ether and ethyl phenylcyanoacetate) both showed large interactions although the size of the interactions would be expected to depend on the structure of the compounds. Included in the test compounds were several compounds with both aliphatic and aromatic substituents (e.g. 2-hydroxybenzyl alcohol), these compounds also showed significant interactions, CRIPES currently contains no information on the interactions of these substituents.

f) Other compounds

It is difficult to define for which compounds the experimental and calculated retention indices can be

regarded as a close fit. The anticipated accuracy of an individual retention index value was determined as +/- 10 units (Chapter 3) but this uncertainty could therefore apply to each substituent and interaction value. The closeness of the fit between the experimental and calculated retention indices would therefore be expected to be a function of the number of substituents. A larger uncertainty was also anticipated with low retention indices and at 80% organic modifier.

Some compounds, e.g thymol, 2,5-dimethylphenol and 3,4-dimethoxyacetophenone, had calculated retention indices which were within +/- 10 (for each substituent) of the experimental values. However, for many compounds there were quite large differences between the experimental and calculated retention indices showing that the interactions are not accounted for by the interaction indices included in the database.

Clearly the present method of accounting for the interactions between substituents is not sufficient to enable the accurate calculation of retention indices. However, the results with some compounds suggest that this approach could provide a useful method of calculating retention. A comparison of the capacity factors suggested that in most cases a reasonable indication of the retention would be given. It should be possible to use the system to give separation conditions for retention within a reasonable time.

Although these compounds were examined as test compounds, in many cases together with suitably selected model compounds these polyfunctional compounds could be in used in future to to extend and refine the database of

interaction terms.

8.8 COMPARISON OF A COMMERCIAL RETENTION PREDICTION SYSTEM, DRYLAB I AND CRIPES

A copy of the commercial prediction system, Drylab I, was supplied by LC Resources Inc., USA. The aim of this section was to briefly examine the use of this system and to compare the determine how the results compared to those obtained using CRIPES.

Drylab I is described in the manual[216] as an expert system to simulate chromatographic separations, however, it does not appear to fulfil the criteria frequently quoted for expert systems. Its operation has been described extensively by Snyder and co-workers in a number of publications[167-169,217-220].

The user is required to input the retention times of the compounds using the information from two isocratic separations (using different proportions of the same modifier) or a single gradient elution run. Using this information the program should be able to calculate the resolution at fixed intervals and therefore the optimum separation conditions. Although the user is able to display a graph of the change in resolution with eluent composition, the program did not actually select or state the optimum separation conditions.

The user is not required to input a column void volume value but the program calculated a value using an expression describing the volume of the stationary phase and the

efficiency of the column. It is possible to recalculate the resolution using an efficiency value entered by the user. A linear relationship between the log capacity factor and mobile phase composition is assumed. However, as shown earlier in this study (Chapter 4) for many compounds this assumption is not valid and could lead to considerable problems with the extrapolation of data. The resolution in Drylab is calculated using an equation similar to that employed by CRIPES.

The results obtained from Drylab and CRIPES were compared for the simulated separation of three sets of two compounds, benzyl alcohol and benzamide, 2-phenylethanol and benzyl chloride, and 3-phenyl-1-propyl bromide and 3-phenyl-1-propyl chloride. Experimental retention times from the present study of these compounds (Table 8.10) were entered into Drylab as the values for two different isocratic eluents. It was not possible to input the retention times at 80% methanol for benzamide and 80% acetonitrile for 2-phenylethanol as the retention times (1.067 and 1.104 minutes) were smaller than the calculated column void volume.

Using these values Drylab calculated capacity factors for intermediate and extrapolated eluent compositions. The capacity factors (Table 8.12) were different to those obtained experimentally (Table 8.11) due to the use of a different column void volume value. The results were very dependent on the initial values that were entered. Large differences were found if the if range was extrapolated to 30% from 50% for 2-phenylethyl bromide (46.32 from 30% acetonitrile and 24.02 based on the retention times at 50% acetonitrile) and 2-phenylethyl chloride (35.74 based on 30%

Table 8.10: Retention times of compounds used to test Drylab and CRIPES

Compound	Modifier	Retention time Organic Modifier Proportion (%)					
		30	40	50	60	70	80
benzamide	MeOH	--	2.448	1.437	1.254	1.120	1.067
benzyl alcohol	MeOH	--	3.230	2.346	1.592	1.425	1.167
2-phenylethanol	MeOH	--	4.988	3.483	1.939	1.584	1.236
benzyl cyanide	MeOH	--	5.650	3.336	1.967	1.575	1.226
2-phenylethanol	MeCN	3.146	2.040	1.532	1.302	1.173	1.104
benzyl cyanide	MeCN	6.678	3.590	2.236	1.649	1.340	1.175
2-phenylethyl bromide	MeCN	52.696	17.289	7.019	2.975	2.401	1.724
2-phenylethyl chloride	MeCN	40.908	14.130	6.003	3.313	2.194	1.614

Table 8.11: Experimental capacity factors of compounds used to test drylab and CRIPES

Compound	Modifier	capacity factor Organic Modifier Proportion (%)					
		30	40	50	60	70	80
benzamide	MeOH	-	1.18	0.68	0.42	0.33	0.25
benzyl alcohol	MeOH	-	2.31	1.31	0.84	0.56	0.41
2-phenylethanol	MeOH	-	4.11	2.42	1.28	0.74	0.50
benzyl cyanide	MeOH	-	4.79	2.28	1.31	0.73	0.48
2-phenylethanol	MeCN	2.71	1.56	1.00	0.68	0.53	0.43
benzyl cyanide	MeCN	6.87	3.50	1.92	1.13	0.75	0.52
2-phenylethyl bromide	MeCN	61.35	20.70	8.18	3.83	2.13	1.23
2-phenylethyl chloride	MeCN	47.24	16.74	6.85	3.28	1.86	1.09

acetonitrile and 20.07 from 50% acetonitrile). Although Drylab puts up a warning when the extrapolation is excessive neither of these examples produced the warning. The capacity factors predicted by Drylab decrease very rapidly with increasing organic modifier proportion of the eluent, this was probably also due to the large column void volume value used. In all cases capacity factors were different to the experimental k' values (Table 8.11). The results were also compared to those calculated by CRIPES (Table 8.13), in this case the calculated k' were closer to the experimental values. The resolutions obtained by the different prediction methods were of a similar magnitude although they were not the same and would probably result in different optima.

Overall CRIPES appears to be useful where the molecular structure of the analytes was known as the calculation of capacity factor and resolution can be done without any prior experimental work. However if the molecular structure was not known an approach similar to that of Drylab would have to be used, experimental data on separations of the sample would be required and the accuracy of the predicted values depends very much on which experimental values were input.

Table 8.12: Capacity factors and resolution calculated using Drylab using retention times in Table 8.8

Eluents for which retention times were entered	Eluent (%)	Calculated capacity factor k'_1 benzamide	k'_2 benzyl alcohol	R_s
40 & 70 MeOH	[40]	1.20	1.90	6.89
	50	0.20	1.00	12.49
	60	0.03	0.53	9.66
	[70]	0.01	0.28	5.99
	80	0.00	0.15	3.42
40 & 60 MeOH	[40]	1.20	1.90	6.89
	50	0.39	0.90	7.82
	[60]	0.13	0.43	5.94
	70	0.04	0.20	3.64
	*80	0.01	0.10	1.99
50 & 70 MeOH	*40	2.06	2.20	1.13
	[50]	0.29	1.11	12.02
	60	0.04	0.56	9.92
	[70]	0.01	0.28	5.99
	80	0.00	0.14	3.26
		2-phenylethanol	benzyl cyanide	
30 & 70 MeCN	[30]	1.83	5.00	17.98
	40	0.75	2.24	14.90
	50	0.31	1.01	10.48
	60	0.13	0.45	6.27
	[70]	0.05	0.20	3.32
	80	0.02	0.09	1.64
40 & 60 MeCN	30	1.24	4.78	22.10
	[40]	0.83	2.22	13.77
	50	0.56	1.03	6.61
	[60]	0.38	0.48	1.84
	70	0.22	0.25	0.59
	*80	0.17	0.10	1.45
		3-phenylethyl bromide	3-phenylethyl chloride	
30 & 80 MeCN	[30]	46.32	35.74	6.30
	40	19.07	14.89	5.81
	50	7.85	1.27	5.12
	60	3.23	2.59	4.13
	70	1.33	1.08	2.87
	[80]	0.55	0.45	1.65
50 & 80 MeCN	30	24.02	20.07	4.28
	40	11.28	9.39	4.17
	[50]	5.30	4.39	3.87
	60	2.49	2.05	3.31
	70	1.22	1.17	2.50
	[80]	0.55	0.45	1.65

* extrapolated value warning that may not be reliable

Table 8.13: Retention indices, capacity factors and resolution calculated by CRIPES

Eluent Composition (%)	Compound k'_1	k'_2	RI	RI	R_s
	benzamide	benzylalcohol	benzamide	benzylalcohol	
40 MeOH	1.36	2.89	604	691	9.85
50	0.72	1.54	593	697	8.09
60	0.43	0.86	580	695	5.77
70	0.30	0.53	566	685	3.70
80	0.24	0.36	551	667	2.20
	2-phenyl ethanol	benzyl cyanide	2-phenyl ethanol	benzyl cyanide	
30 MeCN	2.57	6.52	711	829	13.13
40	1.50	3.38	693	821	10.74
50	0.97	1.84	682	807	7.63
60	0.68	1.06	678	787	4.63
70	0.50	0.65	681	763	2.31
80	0.38	0.42	690	733	0.81
	2-phenylethyl bromide	2-phenylethyl chloride	2-phenylethyl bromide	2-phenylethyl chloride	
30 MeCN	54.86	41.90	1101	1067	5.80
40	20.54	16.41	1104	1069	4.79
50	8.38	6.93	1104	1067	3.85
60	3.77	3.21	1100	1060	2.95
70	1.89	1.64	1092	1048	2.15
80	1.06	0.94	1079	1031	1.55

CHAPTER 9

EXTENDING THE PREDICTION SYSTEM USING PUBLISHED DATA

The database accessed by CRIPES can be extended using purely experimental data, however, it may be also possible to make use of published data to reduce the amount of additional experimental work.

Substituent indices based on retention indices should be highly reproducible and the values should be transferable between columns of the same packing type and make of material. The work described in Chapter 3 showed that for a single batch of one packing material the retention indices were very reproducible, although the capacity factors differed between columns. Previous work has shown that retention indices were sensitive to selectivity changes between columns containing different brands of nominally equivalent packings (i.e. ODS-silicas) but the variations were considerably less than observed with capacity factors[142]. It should therefore be possible to use the values derived in the present work to calculate retention indices for separations on other ODS-silica columns. Conversely it should be possible to use retentions reported on other ODS columns to extend the database.

There are several possible sources of data which could be used to expand the prediction method by either providing additional information on substituent indices or interactions between substituents. The most immediately accessible data is that expressed using the retention index scale, for example previous studies done in this laboratory

on retention indices of test compounds and drugs[134,144,145,148,149]. The use of this data may present problems for direct comparison with the current work as none of the papers have included benzene, the parent used in this study, which meant that it is not possible to directly compare the substituent indices. However it is possible to compare the retention indices in both methanol[134] and acetonitrile[145] which have been measured across eluent ranges comparable to those in this study.

Secondly there are several compilations of capacity factors[63,65,93,221,222] which included at least two of the alkylarylketones and therefore enable retention indices to be calculated by interpolation or extrapolation. However, in these papers there was no indication as to whether the retentions were measured on a single day which may lead to uncertainties in the data.

Finally work by Smith[142] has shown that the retention index of methyl benzoate was virtually constant on different ODS columns and across the eluent ranges, mean *RI* = 905 in methanol-water eluents and mean *RI* = 886 in acetonitrile-water eluents. It could therefore be used as a secondary retention index standard to estimate retention indices, provided that one of the alkylarylketones was present in the set of data. Methyl benzoate and acetophenone have frequently been included in compilations of data[28,29,97,223-226] but because the retention indices are so close the use of this two point line may require quite large extrapolations resulting in an uncertainty in the calculated retention index values.

Two other sources of data could be used to pin-point compounds where significant interactions would occur. The

first of these was the group contribution papers described in Chapter 1, although the data contained in these papers could probably not be used for direct calculation of either substituent indices or interaction indices. By identifying compounds in which the calculated retention differed considerably from the experimental retention it should be possible to identify compounds or substituent combinations for further study. The further possible use of published data would be the use of octanol-water partition coefficient values. As described in a previous chapter (Chapter 7) there should be a relationship between the interaction increments observed for calculations of log *P* by addition of π values and the interactions between substituent in RP-HPLC. The relationship is probably not sufficiently linear to allow the direct use of the relationship to calculate, RP-HPLC interactions, however it should be possible to use the data to identify substituent combinations for further study.

9.1 COMPARISON OF RETENTION INDICES DETERMINED PREVIOUSLY WITH VALUES CALCULATED IN THIS STUDY

The retention indices of a number of test compounds have been determined on several columns[142,145,148] (Table 9.1). The retention indices were measured at single eluent compositions. With most compounds there is quite good agreement between the retention indices on the different columns, the largest differences being observed with the most polar compounds 2-phenylethanol, benzyl alcohol and *p*-cresol. These compounds were more sensitive to the changes in the stationary phase in previous studies by Smith[142] and it has been suggested that they could be used to compare

Table 9.1: Retention indices determined at a single eluent composition on different columns

Compound	Retention index Column[a]								
	ODS-H3	ODS-H5	ODS-T	ODS-S	ODS-Z	ODS-P	ODS-L	ODS-H5	ODS-S2
methanol-buffer 40:60									
	From reference [148]								This work
2-phenylethanol	776	778	753	736	719	751	-	-	756
nitrobenzene	828	828	838	823	818	823	-	-	851
p-cresol	796	798	769	745	726	752	-	-	775
toluene	989	989	985	955	976	936	-	-	983
methanol-water 70:30									
	From reference [142]								This work
2-phenylethanol	779	783	754	743	702	767	758	-	741
nitrobenzene	857	862	875	838	843	849	843	-	874
p-cresol	794	810	776	759	703	753	777	-	748
toluene	1055	1062	1046	988	1040	965	1055	-	1063
methyl benzoate	910	909	902	895	906	902	913	-	910
acetonitrile-water 50:50									
	From reference [142]							ref[145]	This work
2-phenylethanol	694	694	692	697	673	696	691	681	675
nitrobenzene	888	891	880	873	859	870	874	871	871
p-cresol	766	760	751	745	719	720	747	751	741
toluene	1030	1027	1019	1009	1027	984	1026	1021	1036
methyl benzoate	888	881	890	889	887	881	890	886	900

[a] ODS-H3, 3 µm ODS Hypersil; ODS-H5, 5 µm ODS Hypersil; ODS-T, Techsil 5 C-18; ODS-S, Spherisorb ODS; ODS-Z, Zorbax ODS; ODS-P, Partisil 10 ODS; ODS-L, Lichrosorb RP-18; ODS-S2, Spherisorb ODS2;

column selectivity[142]. As the earlier studies were carried out at ambient temperature[142] this may be another source of error resulting in differences in the retention indices between the previous studies and this study. The results suggest that not only could the substituent index equations derived in this study be used for prediction of retention indices on other columns (although obviously with a larger uncertainty in the value) but other retention index data could be used for determining interactions and other substituent indices.

9.2 RETENTION INDICES CALCULATED FROM PUBLISHED DATA

The retention indices can be calculated from sets of data by either using the published capacity factors of two or more alkylarylketones[63,65,93,221] or the capacity factors of acetophenone and methyl benzoate[29,227]. The retention indices calculated using these methods would be expected to have a larger experimental error than those calculated using the full set of alkylarylketones but the method may be useful to estimate retention indices.

Retention indices (Tables 9.2 and 9.3) have been calculated for a range of compounds on different columns and in different eluents. In many cases a single eluent composition has been studied and these enable a comparison of the different columns. For the purposes of this study the results have only been included if benzene was one of the compounds studied, some sources of calculated values have therefore been omitted[29,120,224,225]. In many cases the calculated values were close to the experimental values again the main exceptions were the very polar compounds.

9.3 SUBSTITUENT INCREMENTS USING RETENTION INDICES CALCULATED FROM PUBLISHED CAPACITY FACTORS

The retention indices could be calculated over a range of eluent compositions in two of the papers[69,226]. One of these sets of data has been selected for further study as this paper contained capacity factors measured in three eluent systems (methanol-water, acetonitrile-water and tetrahydrofuran-water)[226]. The retention index of benzene was included in this data set and this has been used to calculate substituent increments. No attempt has been made to calculate a fitted parent index equation but the calculations were carried out using the experimentally derived value (Tables 9.4 and 9.5).

There were differences between these values and those tabulated previously (Chapter 6) but in many cases the differences were quite small suggesting that if the change in the retention index of the parent was known then the substituent indices could be used to calculate retention indices for other columns. The results also suggest that it may be possible to extend the database by using the published data. For example several of the data sets contain information on the retention of naphthalene and its derivatives, these values could be used to calculate the substituent indices for this set of compounds (they would be expected to be very similar to the benzene substituent indices). The change in retention index of naphthalene across the eluent ranges would enable a parent index equation to be calculated and therefore this parent to be included in the database accessed by CRIPES. The retention indices of poly-substituted compounds could also be used to

Table 9.3: Retention indices calculated from published data in acetonitrile eluents

Compound	Retention Index								
								This work	
	Column								
	Nucleosil[226]		Chromosorb LC-7[63]		Hypersil-ODS[227]			Spherisorb ODS2	
	Acetonitrile proportion (%)								
	30	50	30	50	30	30	50	30	50
aniline	-	-	751	833	-	-	-	691	694
anisole	895	915	-	-	895	892	886	900	912
benzaldehyde	788	789	-	-	786	783	779	781	786
benzamide	-	-	601	525	561	547	557	568	521
benzene	893	915	889	913	897	892	880	910	940
benzonitrile	819	818	-	-	819	820	821	814	813
biphenyl	-	1184	-	-	-	-	-	1201	1196
bromobenzene	-	-	--	1039	-	-	-	1041	1065
chlorobenzene	998	1019	998	1010	992	991	986	1014	1037
methyl benzoate	886*	886*	885	890	886*	886*	886*	890	900
nitrobenzene	869	882	864	856	869	872	871	869	871
phenol	699	672	704	673	716	700	695	695	674
toluene	987	1016	989	1010	985	982	972	1005	1036

Compound	Retention Index							
							This work	
	Column							
	Nucleosil[226]		Chromosorb LC-7[63]		Unisil Q C18[63]	FineSIL[28]	Spherisorb ODS2	
	Acetonitrile proportion (%)							
	60	70	60	70	60	65	60	70
aniline	-	-	696	1097	-	751	695	691
anisole	907	904	-	-	-	886	915	917
benzaldehyde	794	796	-	-	-	779	788	788
benzamide	-	-	532	517	557	-	511	509
benzene	907	904	925	907	945	913	951	960
benzonitrile	814	814	-	-	-	805	808	799
biphenyl	1138	1142	-	-	-	-	1195	1194
bromobenzene	-	-	1058	1067	1070	-	1074	1084
chlorobenzene	1005	1011	1028	1017	1038	-	1045	1053
methyl benzoate	886*	886*	888	887	890	886*	894	897
nitrobenzene	869	855	873	837	865	852	864	853
phenol	666	650	658	647	687	620	660	639
toluene	995	1010	1021	1007	1041	999	1046	1054

* defined value used in calculation

Table 9.2: Retention indices calculated from published data in methanol eluents

Compound	Retention index								
								This work	
	Column								
	Nucleosil[226]		Lichrosorb[223]				Alltech[222]	Spherisorb ODS2	
	Methanol proportion (%)								
	50	60	50	50	50	50	55	50	60
aniline	-	-	-	-	-	-	661	658	657
anisole	901	909	892	877	865	887	917	904	904
benzaldehyde	764	774	767	770	687	765	-	777	777
benzamide	-	-	579	583	590	612	568	589	578
benzene	898	911	900	875	857	895	925	915	938
benzonitrile	772	766	788	778	782	772	772	788	775
biphenyl	-	1233	-	-	-	-	1234	1222	1231
bromobenzene	-	-	-	-	-	-	1054	1051	1065
chlorobenzene	1023	1030	1005	986	974	1009	1028	1021	1036
methyl benzoate	905*	905*	905*	905*	905*	905*	913	904	904
nitrobenzene	848	860	835	835	834	839	-	857	864
phenol	675	684	726	696	679	711	-	683	680
toluene	1011	1023	993	977	961	-	1034	1010	1038

Compound	Retention index					
					This work	
	Column					
	Nucleosil[226]		FineSIL[28]	Perkin-Elmer C18[221]	Spherisob ODS2	
	Methanol proportion (%)					
	70	80	75	75	70	80
aniline	-	-	714	645	659	639
anisole	928	924	915	943	910	914
benzaldehyde	771	791	744	761	775	784
benzamide	-	-	-	531	570	551
benzene	931	926	-	946	958	983
benzonitrile	760	774	736	762	774	760
biphenyl	1236	1146	-	-	1247	1270
bromobenzene	-	-	-	1084	1088	1110
chlorobenzene	1037	1000	-	1015	1051	1072
methyl benzoate	905*	905*	905*	920	910	914
nitrobenzene	869	866	855	866	874	874
phenol	662	679	616	662	671	650
toluene	1041	1013	-	1069	1063	1090

* defined value used in calculation

Table 9.4: Aromatic substituent increments calculated for Nucleosil column in methanol

Substituent	Substituent increment									
						This work				
	Column									
	Nucleosil 10-RP18[226]					Spherisorb ODS2				
	Methanol proportion (%)									
	40	50	60	70	80	40	50	60	70	80
OCH_3	13	3	-2	-3	-2	-1	-9	-20	-27	-28
CHO	-108	-134	-137	-160	-135	-111	-136	-161	-186	-198
CN	-99	-126	-145	-171	-152	-109	-133	-163	-187	-222
Ph	-	-	322	305	220	320	309	293	286	288
Cl	-	125	119	106	74	113	108	98	90	90
NO_2	-34	-50	-51	-62	-96	- 34	- 56	- 74	- 87	-108
OH	-192	-223	-227	-269	-247	-205	-230	-258	-290	-332
CH_3	-	113	112	110	87	102	106	101	104	113
$COCH_3$	-69	-98	-111	-131	-126	-80	-107	-135	-158	-178
CO_2CH_3	36	7	-6	25	21	14	-9	-34	-51	-68

Table 9.5: Aromatic substituent increments calculated for a nucleosil column in acetonitrile

Substituent	Substituent increment											
							This work					
	Column											
	Nucleosil 10-RP18[226]						Spherisorb ODS2					
	Acetonitrile proportion (%)											
	30	40	50	60	70	80	30	40	50	60	70	80
OCH_3	2	-10	-9	-12	-20	-15	-10	-19	-28	-36	-41	-50
CHO	-105	-127	-135	-125	-128	-132	-129	-141	-154	-163	-170	-184
CN	-74	-95	-106	-105	-110	-132	-96	-110	-127	-143	-151	-175
Ph	-	268	260	219	218	204	291	271	256	244	236	232
Cl	105	96	95	86	87	87	104	100	97	94	95	95
NO_2	-24	-33	-42	-50	-69	-77	-41	-53	-69	-87	-105	-127
OH	-194	-225	-252	-253	-274	-312	-215	-240	-266	-291	-319	-318
CH_3	94	90	92	76	86	87	95	95	96	94	95	95
CO_2CH_3	-7	-18	-29	-21	-18	-19	-20	-37	-40	-57	-64	-66
$COCH_3$	-93	-104	-115	-107	-104	-105	-110	-129	-142	-153	-159	-160

calculate further interaction indices and to test any generalised theory (see Chapter 7) without extensive further experimental work. The results show that the substituent indices calculated in this work should provide a reasonable estimate of the retention indices on other columns with different brands of stationary phases. There would probably be larger errors than for prediction on the same packing material and the largest differences would be observed when the substituent increments were large (such as OH).

9.4 INTERACTION INCREMENTS USING RETENTION INDICES CALCULATED FROM PUBLISHED CAPACITY FACTORS

The aim in calculating the interaction increments was to examine whether those calculated from experimental data (Chapter 7) were similar to those calculated from retention indices based on published capacity factors. Secondly it might be to possible identify other substituent pairs, not included in the present study, with large interaction. The interaction increments were calculated for one set of data in a single organic modifier system. The results were calculated in acetonitrile-buffer for the results published by Schoenmakers *et al.*[226].

The increments (Table 9.6) were calculated using the benzene and substituent increments given in Table 9.5. Large interaction increments were observed for the substituent pairs $CO_2C_2H_5$ + 2-$CO_2C_2H_5$, and CO_2CH_3 + 2-CO_2CH_3. The value of the retention index of dimethyl phthalate calculated in this work (Chapter 8) also differed considerably from the experimental value, the differences in the Chapter 8 (about

Table 9.6: Interaction increments calculated from retention indices using published capacity factors

Substituents (on benzene)	Interaction increment Acetonitrile proportion (%)					
	30	40	50	60	70	80
COR + Ph	-	-34	-74	-83	-82	-78
CHO + 4-Cl	-57	-87	-	-	11	9
OH + 4-Cl	27	17	14	13	-2	-9
OH + 3-CH_3	-	-8	-6	-18	-22	-37
OH + 2-CH_3	10	-2	-7	-11	-8	1
Cl + 2-Cl	-	-25	-20	-24	-36	-13
$CO_2C_2H_5$ + 2-$CO_2C_2H_5$	-102	-99	-99	-134	-154	-170
CO_2CH_3 + CO_2CH_3	-45	-48	-48	-66	-	-98
NO_2 + 3-NO_2	24	10	14	12	11	-12
NO_2 + 2-NO_2	44	19	11	2	3	-20
NO_2 + 4-NO_2	25	17	19	14	11	-20
CH_3 + 2-NO_2 + 4-NO_2	2	-12	-19	-33	-33	-53
CH_3 + 2-NO_2 + 6-NO_2	3	-9	-	-29	-33	-53
CH_3 + 3-NO_2 + 4-NO_2	20	-6	-17	-30	-41	-53
OR + Ph	-	97	91	89	74	76
OH + 4-CHO	43	24	19	15	-25	-51
OCH_3 + 4-CHO	9	3	0	-8	-17	-13
CH_3 + 4-CHO	-23	-15	76	70	64	71
NO_2 + 4-$COCH_3$	63	59	59	54	53	31
NO_2 + 4-CHO	39	44	38	38	24	7
OH + 3-NO_2	99	88	94	84	74	74
OH + 2-NO_2	157	169	191	196	222	254
OH + 4-NO_2	95	132	68	59	49	62
OH + 4-Ph	-	-11	-5	-10	-23	1
OH + 2-CH_3 + 4-CH_3	-38	-36	-26	-40	-36	-14

60-80 units) were comparable with those observed in this set of data. There are several substituent pairs which were also considered in the disubstituted compounds (Chapter 7). With the exception of the substituent pairs CH_3 + 4-CHO, OH + 3-NO_2, and OH + 4-NO_2, where there was concern that these compounds might be ionised (Chapter 7), the interaction increments were very similar to those obtained in the present study showing that the interactions between substituents have a similar effect on the retention index on different columns.

9.5 USE OF OCTANOL-WATER PARTITION DATA

In the study of substituent indices (Chapters 4 and 6) there was shown to be an approximately linear relationship between the substituent index and the Hansch substituent constant (π). This relationship might enable the π values to be used to estimate a value for a substituent index. However as the linearity was not high the value obtained would be only a rough guide to the true substituent increment.

Another possible application of the octanol-water partition data could be its use to pin-point compounds in which there were large interactions between substituents. For instance if the octanol-water partition coefficient calculated as the $\Sigma\pi$ was very different to the experimental octanol-water partition coefficient this might suggest a substituent pair for further investigation. Although the results in Chapter 7 suggested that the proposed method of correcting π values for the interactions would not be directly applicable to the calculation of interaction

indices, the compounds which significantly deviate would be expected to be similar in both systems. The paper by Leo[51] which derived the interaction terms for log *P* listed the experimental and $\Sigma\pi$ values for 400 compounds, from which it should be possible to extract particular substituent combinations for further study, for example *o*-nitroaniline or *m*-cyanobenzamide where the experimental octanol-water partition coefficients were 3 to 4 times the calculated values.

Conversely octanol-water partition values could also be used to be used to identify substituent combinations in which interactions were not significant. There might be some problems in using the data as some compounds do not show the same behaviour in the two systems, for example substituted phenols.

Another possible source of data would be the many log *P*, log *k'* sets of data. However, as these do not always make it clear whether the log *P* values were the additive values or measured values, the use of this data to estimate interactions might not be possible. These papers could contain data which would highlight compounds in which the interactions were definitely not the same in HPLC as in octanol-water and would therefore need to be studied individually in the HPLC system individually.

No attempt has been made to use the octanol - water data for this purpose so far in the present study.

9.6 USE OF GROUP CONTRIBUTION DATA

A number of papers have attempted to predict the retention of compounds based on a group contribution approach (see Chapter 1, Section 1.2.6). Although none of these papers used the alkylarylketone retention index scale it might be possible to use the calculated retentions to identify interactions although it will probably not be possible to use the data to derive numerical values.

9.7 EXPANDING CRIPES

Using the sources of data discussed above it should be possible to extend the database from which CRIPES extracts the regression coefficients. In particular the database currently holds relatively little information on the interactions. It may be possible to obtain further information without extensive experimental work, although currently it would appear that a considerable number of experimental retention indices will need to be determined at least to confirm the validity of any proposed prediction. The interaction indices currently held in the spreadsheets are based purely on empirical values valid only for the particular pair of substituents. It is important that to improve the usefulness of the prediction system some underlying rules describing the interactions need to be developed. A system similar to that proposed by Leo[52] to account for the interactions between substituents in octanol-water systems may prove a useful method of calculating interactions (Chapter 7).

CHAPTER 10

CONCLUSION

Using values based on the retention indices of benzene and mono-substituted benzenes or alkylbenzenes it has been possible to derive parent index equations and substituent index equations. These quadratic equations relate the change in the parent index and substituent index to the methanol or acetonitrile proportion of the mobile phase. Using these equations it has been possible to calculate the effect, on the retention index, of interactions between substituents in a number of disubstituted compounds. These substituent index equations could also be used with parent compounds other than benzene, for example naphthalene, with a different parent index equation.

The proposed method of calculating retention indices of "unknown" compounds from their molecular structure as the sum of the parent index, the substituent indices and the interaction indices have been shown to be successful in some cases. However, with many compounds there is currently insufficient information to fully account for the interactions between substituents. In its present form the prediction method requires an empirical interaction index equation to for each interaction and assumes that the interactions are additive. Thus there is no judgement about whether one (or more) interactions would actually dominate in a particular molecule to the exclusion of other interactions. This approach to substituent interactions will not be satisfactory in the long term but a method of accounting for interactions based on a set of "ground rules"

needs to be developed and possible approaches have been considered for future work. Any rules for this purpose would have to be based on a recognition of the types of interactions and their relative size with different substituent combinations. To enable the formulation of such rules a considerably larger amount of data would need to be examined than that contained in this study

The expert system program, CRIPES, proved to be a convenient method of extracting the appropriate data from the spreadsheets. This enabled the easy calculation of retention indices based on molecular structure without the user having to "remember" the rules for the interactions. The method by which retention indices were expressed as quadratic equations relating the change in retention index to either methanol or acetonitrile proportion enabled the calculation of resolution between two peaks. Although the maximum resolution calculated by CRIPES may not be the final optimised separation conditions it should be possible to use the prediction to provide a starting point after which a further optimisation procedure could then be applied.

The commercial expert system, Drylab I, was straightforward to use, however, the results obtained were very dependent on the initial data entered. If the molecular structures of the components of a mixture are not known this approach to the calculation of resolution should prove valuable. However, if the molecular structures of the compounds are known an approach similar to that taken using CRIPES should reduce the time required to select the initial separation conditions.

FUTURE WORK

CRIPES contains information on the change in the index values with proportion of methanol and acetonitrile. Similar equations need to be derived for tetrahydrofuran (THF) the third commonly used RP-HPLC solvent .

By using other aqueous phase buffer pHs it would be possible to calculate the substituent index equations for functional groups which could not be studied due to their ionisation, e.g. carboxylic acid groups and basic amines. It should also be possible to check that the substituent index equations are not dependent on the aqueous phase pH provided that the functional group is not ionised.

The main task is to study in more details interactions between substituents to try to discover any underlying rules about their size and significance. It may be possible to study these interactions without extensive further practical work using published data, although, any proposed rules would need to be validated by experimental work.

REFERENCES

1 R. Kaliszan, *Quantitative Structure-Chromatographic Retention Relationships*, Chemical Analysis Vol. 93, Ed. J.D. Winefordner, Wiley, New York, 1987,

2 Cs. Horvath, W. Melander & I. Molnar, *J. Chromatogr.*, 125 (1976) 129.

3 Cs. Horvath, W. Melander & I. Molnar, *Anal. Chem.*, 49 (1977) 142.

4 A. Tchapla, H. Colin & G. Guiochon *Anal. Chem.*, 56 (1984) 621

5 R.B. Sleight, *J. Chromatogr.*, 83 (1973) 31

6 J. Kriz, L. Vodicka, J. Puncocharova & M. Kuras, *J. Chromatogr.*, 219 (1981) 53

7 P. Dufek & E. Smolkova, *J. Chromatogr.*, 257 (1983) 247

8 P. Dufek, *J. Chromatogr.*, 281 (1983) 49

9 P. Dufek, *J. Chromatogr.*, 299 (1984) 109

10 C. Viseras, R. Cela, C.G. Barroso & J.A. Perez-Bustamante, *Anal. Chim. Acta*, 196 (1987) 115

11 J.H. Knox, J. Kriz & E. Adamcova, *J. Chromatogr.*, 447 (1988) 13

12 J.F. Schabron, R.J. Hurtubise & H.F. Silver, *Anal. Chem.*, 50 (1978) 1911

13 L.R. Schronk, R.D. Grisby & A.R. Hanks, *J. Chromatogr. Sci.*, 19 (1981) 490

14 R.V. Golovnya & D.N. Grigoryeva, *HRC CC, J. High Resolut. Chromatogr. Chromatogr. Commun.*, 9 (1986) 584

15 N.I. Sveltova, D.N. Grigoryeva, L.I. Sokolova & R.V. Golovnya, *J. Chromatogr.*, 364 (1986) 203

16 L.B. Kier & L.H. Hall, *Molecular Connectivity in Structure-Activity Analysis*, Research Studies Press

Ltd., Herts., England, 1986

17 L.B. Kier, L.H. Hall, W.J. Murray & M. Randic, *J. Pharm. Sci.*, 64 (1975) 1971

18 I.S. Lurie & A.C. Allen, *J. Chromatogr.*, 292 (1984) 283

19 P. Lehtonen, *J. Chromatogr.*, 398 (1987) 143

20 M.J.M. Wells, C.R. Clark & R.M. Patterson, *J. Chromatogr.*, 235 (1982) 61

21 M.J.M. Wells, C.R. Clark & R.M. Patterson, *Anal. Chem.*, 58 (1986) 1625

22 R.J. Hurtubise, T.W. Allen & H.F. Silver, *J. Chromatogr.*, 235 (1982) 517

23 J. Burda, M. Kuras, J. Kriz & L. Vodicka, *Fresnius Z. Anal. Chem.*, 321 (1985) 349

24 R.M. Smith, *J. Chromatogr.*, 209 (1981) 1

25 D. Noel & P. Vangheluwe, *J. Chromatogr.*, 388 (1987) 75

26 F.P. Burke & H.A. Winschell, *Anal. Chem.*, 59 (1987) 2586

27 N. Funasaki, S. Hada & S. Neya, *J. Chromatogr.*, 361 (1986) 33

28 K. Jinno & K. Kawasaki, *Chromatographia*, 18 (1984) 90

29 K. Jinno & K. Kawasaki, *J. Chromatogr.*, 316 (1984) 1

30 K. Jinno & M. Kuwajima, *Chromatographia*, 21 (1986) 541

31 N. Adler, D. Babic & N. Trinajstic, *Fresnius Z. Anal. Chem.*, 322 (1985) 426

32 K. Jinno, T. Nagoshi, N. Tanaka, M. Okamoto, J.C. Fetzer & W.R. Biggs, *J. Chromatogr.*, 386 (1987) 123

33 K. Jinno & K. Kuwajima, *Chromatographia*, 22 (1986) 13

34 K. Jinno & K. Kawasaki, *Chromatographia*, 17 (1983) 445

35 R.H. Rohrbaugh & P.C. Jurs, *Anal. Chem.*, 59 (1987) 1048

36 S.A. Wise, W.J. Bonnett, F.R. Guenther & W.E. May, *J. Chromatogr. Sci.*, 19 (1981) 457

37 J.F. Schabron, R.J. Hurtubise & H.F. Silver, *Anal. Chem.*, 49 (1977) 2253

38 M.N. Hasan & P.C. Jurs, *Anal. Chem.*, 55 (1983) 263

39 T. Hanai & J. Hubert, *J. Chromatogr.*, 290 (1984) 197

40 T. Hanai & J. Hubert, *J. Chromatogr.*, 302 (1984) 89

41 T. Hanai, *J. Chromatogr.*, 332 (1985) 189

42 Feng Yugi, Zhu Pengling & Hu Zhide, *Chromatographia*, 25 (1988) 382

43 M. D'Amboise & M.J. Bertrand, *J. Chromatogr.*, 361 (1986) 13

44 H.J. Mockel, G. Welter & H. Melzer, *J. Chromatogr.*, 388 (1987) 255

45 H.J. Mockel, H. Hofler & H. Melzer, *J. Chromatogr.*, 388 (1987) 267

46 H.J.Mockel, H. Hofler & H. Melzer, ,*J. Chromatogr.* 388 (1987) 275

47 H.J. Mockel, F. Hofler and H. Melzer, *J. Chromatogr.*, 388 (1987) 285

48 C. Hansch & A. Leo, *Substituent constants for correlation analysis in chemistry and biology*, Wiley, New York, 1979

49 T. Fujito, *J. Pharm. Sci.*, 72 (1983) 285

50 T. Fujita, *Prog. Phys. Org. Chem.*, 14 (1983) 75

51 A. Leo, *J. Chem. Soc. Perkin Trans. 2*, (1983) 825

52 A.J. Leo, *Partition Coefficient Determination and Estimation*, Ed. W.J. Dunn III, J.H. Block and R.S. Pearlman, Pergamon Press Inc., New York, 1986

53 R.F. Rekker, *The Hydrophobic fragmental constant*, Elsevier Scientific Publishing Co., Amsterdam, 1977

54 T. Braumann, *J. Chromatogr.*, (1986), 373,

55 R. Kaliszan, *CRC Crit. Rev. Anal. Chem.*, 16 (1986) 323

56 K. Miyake, N. Mizuno & H. Terada, *Chem. Pharm. Bull.*, 34 (1986) 4787

57 J.J. Sabatka, D.J. Minick, T.K. Shumaker, G.L. Hodgson Jr. & D.A. Brent, *J. Chromatogr.*, 384 (1987) 349

58 Th. Braumann, H.-G. Geneiser, C. Lullman & B. Jastorff, *Chromatographia*, 24 (1987) 777

59 P.J. Schoenmakers, H.A. Billiet & L. De Galan, *J. Chromatogr.*, 185 (1979) 179

60 P.J. Schoenmakers, H.A. Billiet & L. De Galan, *J. Chromatogr.*, 282 (1983) 107

61 T. Hanai & J. Hubert, *J. Liq. Chromatogr.*, 8 (1985) 2463

62 T. Hanai & J. Hubert, *HRC CC, J. High Resolut. Chromatogr. Chromatogr. Commun.*, 6 (1983) 20

63 T. Hanai, C. Tran & J. Hubert, *HRC CC, J. High Resolut. Chromatogr. Chromatogr. Commun.*, 4 (1981) 454

64 T. Hanai, K.C. Tran & J. Hubert, *J. Chromatogr.*, 239 (1982) 385

65 T. Hanai & J. Hubert, *J. Chromatogr.*, 239 (1982) 527

66 T. Hanai & J. Hubert, *HRC CC, J. High Resolut. Chromatogr. Chromatogr. Commun.*, 7 (1984) 524

67 A. Shalaby, Zs. Budvari-Brany & Gy. Szasz, *J. Liq. Chromatogr.*, 7 (1984) 1133

68 A. Shalaby, Zs. Budvari-Barany, Gy. Szasz & I. Hermecz, *Acta Pharm. Hung.*, 55 (1985) 249

69 J.K. Baker, *Anal. Chem.*, 51 (1979) 1693

70 J.K. Baker, G.J. Hite, M. Reamer & P. Salva, *Anal. Chem.*, 56 (1984) 2932

71 J.K. Baker, R.E. Skelton, T.N. Riley & J.R. Bagley, *J. Chromatogr. Sci.*, 18 (1980) 153

72 J.K. Baker, *J. Liq. Chromatogr.*, 4 (1981) 271

73 J.K. Baker & E.K. Fifer, *J. Pharm. Sci.*, 69 (1980) 590

74 K. Jinno & K. Kawasaki, *Anal. Chim. Acta*, 152 (1983) 25

75 K. Jinno & K. Kawasaki, *Chromatographia*, 17 (1983) 337

76 K. Jinno, *Anal. Lett.*, 15(A19) (1982) 1533

77 K. Jinno & K. Kawasaki, *J. Chromatogr.*, 298 (1984) 326

78 K. Jinno & M. Kuwajima, *LC-GC*, 5 (1987) 1038

79 S.F.Y. Li & H.K. Lee, *Chromatographia*, 25 (1988) 515

80 T.L. Hafkenscheid & E. Tomlinson, *J. Chromatogr.*, 218 (1981) 409

81 T.L. Hafkenscheid & E. Tomlinson, *Int. J. Pharm.*, 17 (1983) 1

82 E. Tomlinson & T.L. Hafkenscheid, *Partition Coefficient Determination and Estimation*, Ed. W.J. Dunn III, J.H. Block and R.S. Pearlman, Pergamon Press Inc., New York, 1986

83 T.L. Hafkenscheid & E. Tomlinson, *J. Chromatogr.*, 264 (1983) 47

84 R. Tijssen, H.A.H. Billiet & P.J. Schoenmakers, *J. Chromatogr.*, 122 (1976) 185

85 K. Jinno, *Chromatographia*, 20 (1985) 743

86 M.J. Kamlet, R.M. Doherty, J-L.M. Abboud, M.H. Abraham & R.W. Taft, *ChemTech*, (1986) 566

87 J.G. Dorsey & B.P. Johnson, *Chim. Oggi*, (1986) 23

88 B.P. Johnson, M.G. Khaledi & J.G. Dorsey, *J. Chromatogr.*, 384 (1987) 221

89 B.P. Johnson, M.G. Khaledi & J.G. Dorsey, *Anal. Chem.*,

58 (1986) 2354

90 M.J. Kamlet, R.M. Doherty, P.W. Carr, D. Mackay, M.H. Abraham & R.W. Taft, *Environ. Sci. Techol.*, 22 (1988) 503

91 P.W. Carr, R.M. Doherty, M.J. Kamlet, R. Taft, W. Melander & Cs. Horvath, *Anal. Chem.*, 58 (1986) 2674

92 P.C. Sadek, P.W. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft & M.H. Abraham, *Anal. Chem.*, 57 (1985) 2971

93 J.H. Park, P.W. Carr, M.H. Abraham, R.W. Taft, R.M. Doherty & M.J. Kamlet, *Chromatographia*, 25 (1988) 373

94 I. Molnar & C. Horvath, *J. Chromatogr.*, 145 (1978) 371

95 P.C. Sadek & P.W. Carr, *J. Chromatogr.*, 288 (1984) 25

96 A. Tsantila-Kakoulidou, N. El Tayar, H. van de Waterbeemd & B. Testa, *J. Chromatogr.*, 389 (1987) 33

97 M.C. Spanjer & C.L. De Ligny, *Chromatographia*, 20 (1985) 120

98 M.C. Spanjer & C.L. De Ligny, *J. Chromatogr.*, 253 (1982) 23

99 E. Tomlinson, H. Poppe & J.C. Kraak, *Int. J. Pharm.*, 7 (1981) 225

100 K. Glowniak & M.L. Bieganowska, *J. Chromatogr.*, 370 (1986) 281

101 B-K. Chen & C. Horvath, *J. Chromatogr.*, 171 (1979) 15

102 R. Gill, S.P. Alexander & A.C. Moffat, *J. Chromatogr.*, 218 (1981) 639

103 S.P. Assenza & P.R. Brown, *J. Chromatogr.*, 282 (1983) 477

104 J. Borda, V. Szabo & J. Kelemen, *J. Chromatogr.*, 286 (1984) 113

105 V.Ya. Davydov, *J. Chromatogr.*, 365 (1986) 123

106 J.-X. Huang, E.S.P Bouvier, J.D. Stuart, W.R. Melander

& Cs. Horvath, *J. Chromatogr.*, 330 (1985) 138

107 A. Bylina, L. Gluzinski, K. Lesniak, B. Radwanski & P.A. Penczek, *Chromatographia*, 17 (1983) 132

108 T. Braumann, G. Weber & L.H. Grimme, *J. Chromatogr.*, 261 (1981) 329

109 T. Braumann & B. Jastorff, *J. Chromatogr.*, 350 (1986) 105

110 M. Popl, V. Dolansky & J. Fahnrich, *J. Chromatogr.*, 148 (1978) 195

111 F. Smejkal, M. Popl, A. Cihova & M. Zazvorkova, *J. Chromatogr.*, 197 (1980) 147

112 M. Popl, I. Vit & F. Smejkal, *J. Chromatogr.*, 213 (1981) 363

113 H. Magg & K. Ballschmiter, *J. Chromatogr.*, 331 (1985) 245

114 F. Morishita, H, Kakihana & T. Kojima, *Anal. Lett.*, 17(A20) (1984) 2385

115 H. Colin, G. Guiochon & P. Jandera, *Anal. Chem.*, 55 (1983) 442

116 H. Colin, G. Guiochon & P. Jandera, *Chromatographia*, 17 (1987) 83

117 P. Jandera, *J. Chromatogr.*, 314 (1984) 13

118 P. Jandera, *J. Chromatogr.*, 352 (1986) 91

119 P. Jandera, *J. Chromatogr.*, 352 (1986) 111

120 P. Jandera, *Chromatographia*, 19 (1984) 101

121 P. Jandera & M. Spacek, *J. Chromatogr.*, 366 (1986) 107

122 J.D. Baty & S. Sharp, *J. Chromatogr.*, 437 (1988) 13

123 H.A. Cooper & R.J. Hurtubise, *J. Chromatogr.*, 360 (1986) 313

124 E. Marengo, M.C. Gennaro, C. Baiocchi and P. Bertolo, *Chromatographia*, 25 (1988) 413

125 M. Jaroniec & D.E. Martire, *J. Chromatogr.*, 387 (1987) 55

126 M. Jaroniec & D.E. Martire, *J. Liq. Chromatogr.*, 10 (1987) 541

127 S.M. Petrovic, S. Lomic & I. Sefer, *J. Chromatogr.*, 348 (1985) 49

128 H. Lamparczyk & A. Radecki, *Chromatographia*, 18 (1984) 615

129 H. Lamparczyk, *Chromatographia*, 20 (1985) 283

130 H. Lamparczyk, M. Atomura & K. Jinno, *Chromatographia*, 23 (1987) 752

131 E. Kovats, *Helv. Chim. Acta*, 41 (1958) 1915

132 R.M. Smith, *Adv. Chromatogr.*, 26 (1987) 277

133 J.K. Baker & C.-Y. Ma, *J. Chromatogr.*, 169 (1979) 107

134 R.M. Smith, *J. Chromatogr.*, 236 (1982) 313

135 M. Bogusz & R. Aderjan, *J. Chromatogr.* 435 (1988) 43

136 M. Bogusz, R. Aderjan, J.P. Franke & R.A. de Zeeuw, *Z. Rechtsmedizin*, 98 (1987) 263

137 M. Bogusz, *J. Chromatogr.*, 387 (1987) 404

138 M. Bogusz and R. Aderjan, *J. Anal. Toxicol.*, 12 (1988) 67

139 D.W. Hill & K.J. Langner, ,*J. Liq. Chromatogr.*, 10 (1987) 377

140 J.C. Frisvad & U. Thrane, *J. Chromatogr.*, 404 (1987) 195

141 N. Dimov, *Anal. Chim. Acta*, 201 (1987) 217

142 R.M. Smith, *Anal. Chem.*, 56 (1984) 256

143 E. Grushka, H. Colin & G. Guiochon, *J. Chromatogr.*, 248 (1982) 325

144 R.M. Smith, T.G. Hurdley, R. Gill & A.C. Moffat, *Chromatographia*, 19 (1984) 401

145 R.M. Smith, G.A. Murilla & C.M. Burr, *J. Chromatogr.*, 388 (1987) 37

146 R.M. Smith & D.R. Garside, *J. Chromatogr.*, 407 (1987) 19

147 R.M. Smith, T.G. Hurdley, R. Gill & A.C. Moffat, *LC-GC*, 4 (1986) 314

148 R.M. Smith, T.G. Hurdley, R. Gill & A.C. Moffat, *Chromatographia*, 19 (1984) 407

149 R.M. Smith, T.G. Hurdley, R. Gill & A.C. Moffat, *J. Chromatogr.*, 351 (1986) 259

150 R.J. Smith, C.S. Nieass & M.S. Wainwright, *J. Liq. Chromtogr.*, 9 (1986) 1387

151 R.A. Djerki & R.J. Laub, *J. Liq. Chromatogr.*, 10 (1987) 1749

152 R.A. Djerki & R.J. Laub, *J. Liq. Chromtogr.*, 11 (1988) 327

153 M.C. Hennion & R. Rosset, *Chromatographia*, 25 (1988) 43

154 M.J.M. Wells & C.R. Clark, *Anal. Chem.*, 53 1981 1341

155 K. Jinno, N. Ozaki & T. Sato, *Chromatographia*, 17 (1983) 3441

156 P.J.M. van Tulder, J.P. Franke, R.A. de Zeeuw, *HRC CC, J. High Resolut. Chromatogr. Chromatogr. Commun.*, 10 (1987) 191

157 R.J. Laub & S.J. Maddon, *J. Liq. Chromtogr.*, 8 (1985) 173

158 H. Colin, A. Krstulovic, G. Guiochon & J.P. Bounine, *Chromatographia*, 17 (1983) 209

159 M.S. Wainwright, C.S. Nieass, J.K. Haken & R.P. Chaplin, *J. Chromatogr.*, 321 (1985) 287

160 H.J. Mockel & T. Freyholdt, *Chromatographia*, 17 (1983) 215

161 G.E. Berendsen, P.J. Schoenmakers, L. de Galan, G. Vigh, Z. Vargha-Puchony & J. Inczedy, *J. Liq. Chromatogr.*, 3 (1980) 1669

162 H. Colin, A.M. Krstulovic, M-F. Gonnord, G. Guiochon, Z. Yun & P. Jandera, *Chromatographia*, 17 (1983) 9

163 *Lange's Handbook of Chemistry*, 13th Edition, Ed. J.A. Dean, McGraw-Hill Inc., U.S.A. 1985

164 G.B. Cox & R.W. Stout, *J. Chromatogr.*, 384 (1987) 315

165 *CRC Handbook of Chemistry and Physics* 68th Edition 1987-1988, Ed. R.C. Weast, CRC Press Inc., Florida 1987

166 H. Colin, G. Guiochon, Z. Yun, J.C. Diez-Masa & P. Jandera, *J. Chromatogr. Sci.*, 21 (1983) 179

167 L.R. Snyder, J.W. Dolan & M.A. Quarry, *TrAC, Trends Anal. Chem.*, 6 (1987) 106

168 L.R. Snyder & M.A. Quarry, *J. Liq. Chromatogr.*, 10 (1987) 1789

169 M.A. Quarry, R.L. Grob & L.R. Snyder, *Anal. Chem.*, 58 (1986) 901

170 E.D. Katz, K. Ogan & R.P.W. Scott, *J. Chromatogr.*, 352 (1986) 67

171 L.R. Snyder, *J. Chromatogr. Sci.*, 16 (1978) 223

172 H.J.C.F. Nelis & A.P. de Leenheer, *J. Chromatogr.*, 195 (1980) 35

173 R. Gill, S.P. Alexander & A.C. Moffat, *J. Chromatogr.*, 247 (1982) 39

174 H. Colin & G. Guiochon, *J. Chromatogr. Sci.*, 18 (1980) 54

175 R. Kaliszan, *J. Chromatogr.*, 220 (1981) 71

176 K. Wakita, M. Yoshimoto, S. Miyamoto & H. Watanabe, *Chem. Pharm. Bull.*, 34 (1986) 4663

177 R.T. Morrison & R.N. Boyd, *Organic Chemistry*, 3rd

Edition, Allyn and Bacon Inc., U.S., 1973

178 T. Fujita, J. Isawa and C. Hansch, *J. Am. Chem. Soc.*, 86 (1964) 5177

179 G.B. Barlin and D.D. Perrin, *Quarterly Reviews*, 20 (1966) 75

180 Y. Baba, *JAPC*, 7 (1985) 119

181 E. Grushka & I. Atamna, *J. Chromatogr.*, 400 (1987) 5

182 G.A. Taylor, *J. Chromatogr.*, 195 (1987) 375

183 J.C. Berridge, *Techniques for the Automated Optimisation of HPLC Separations*, Wiley, Chichester, 1985

184 G. D'Agnostino & L. Castagnetta, *J. Chromatogr.*, 338 (1985) 1

185 P. Jackson, *Introduction to Expert Systems*, International Computer Science Series, Addison-Wesley Publishing Company, 1986

186 I. Graham and P. Llewellyn-Jones, *Expert Systems Knowledge, Uncertainty and Decision*, Chapman-Hall Computing, London, 1988

187 P.B. Ayscough, S.J. Chinnick, R. Dybowski & P. Edwards, *Chem. Ind.*, 15 (1987) 515

188 T.P. Bridges, M.H. Williams & A.F. Fell, *Chem. Brit.*, 23 (1987) 1085

189 M-P Derde, L. Buydens, C. Guns, D.L. Massart & P.K. Hopke, *Anal. Chem.*, 59 (1987) 1868

190 J.L. Glajch, *LC-GC* 6 (1988) 30

191 J.W. Frazer, *Mikrochimica Acta*, II (1986) 163

192 G.J. Kleywegt, *Lab. Microcomput.*, 6 (1987) 74

193 *Artificial Intelligence in Chemistry*, ACS Symposium Series 306, Ed. T.H. Pierce & B.A. Hohne, ACS, Washington, 1986

194 M. Pownall, *Int. Lab.*, 18 (1988) 6

195 R.E. Dessy, *Anal. Chem.*, 56 (1984) 1201A

196 R.E. Dessy, *Anal. Chem.*, 56 (1984) 1312A

197 S.S. William, J.-L. Excoffier, J.F. Karnicky & S.R. Abbott, *11th International Symposium on Column Liquid Chromatography*, Amsterdam, 1988, Abstract No. F-P8,

198 R. Bach, J. Karnicky & S. Abbott Ch. 22 in, *Artificial Intelligence Applications in Chemistry*, Ed T.H. Pierce & B.A. Hohne A.C.S. (1986), U.S.A.

199 Lu Peichang & Lu Xiamong, *J. Chromatogr.*, 292 (1984) 169

200 Lin Binsheng, Li Haochun, Lu Peichang *LC-GC*, 4 (1986) 1206

201 Zou Hanfa, Chen Long, Zhang Yukui & Lu Peichang *17th International Symposium on Chromatography*, Vienna, 1988 Abstract No. Tue-B L5

202 M.A. Tischler & E.A. Fox, *Comput. Chem.*, 11 (1987) 235

203 H. Gunasingham, B. Srinivasan & A.L. Ananda, *Anal. Chim. Acta*, 182 (1986) 193

204 A.L. Ananda, S.M. Foo and H. Gunasingham, *J. Chem. Inf. Comput. Sci.*, 28 (1988) 82

205 T.P. Bridges, M.H. Williams & A.F. Fell, *Anal. Proc.*, 25 (1988) 43

206 T.P. Bridges, M.H.W. Williams, G.G.R. Seaton and A.F. Fell, *Chromatographia*, 24 (1987) 691

207 G. Musch, M. De Smet & D.L. Massart, *J. Chromatogr.*, 348 (1985) 97

208 G. Musch & D.L. Massart, *J. Chromatogr.*, 370 (1986) 1

209 M. De Smet and D.L. Massart, *TrAC, Trends Anal. Chem.*, 6 (1987) 266

210 P.J. Schoenmakers & M. Mulholland, *Chromatographia*, 25

(1988) 737

211 D. Goulder, T. Blaffert, A. Blokland, L. Buydens, A. Chhabra, A. Cleland, N. Dunand, H. Hindriks, G. Kateman, H. van Leeuwen, D. Massart, M. Mullholland, G. Musch, P. Naish, A. Peeters, G. Postma, P. Schoenmakers, M. de Smet & B. Vandeginste, *17th International Symposium on Chromatography*, Vienna, 1988 Abstract No. Tue-B L9

212 R. Hindriks, L. Butgarde, A. Peeters, M. de Smet, L. Massart, A. Blokland & J. Vink, *17th International Symposium on Chromatography*, Vienna, 1988 Abstract No. I P87

213 L. Massart, L. Buydens, A. Peeters & P.J. Schoenmakers, *17th International Symposium on Chromatography*, Vienna, 1988 Abstract No. I P88

214 M. Mulholland, H. van Leeuwen and B. Vandeginste, *17th International Symposium on Chromatography*, Vienna, 1988 Abstract No. I P89

215 P.J. Schoenmakers, G. Musch, A. Cleland, N. Dunand, P.J. Naish & P.K. de Bokx, *17th International Symposium on Chromatography*, Vienna, 1988 Abstract No. I P93

216 *Drylab I Users Manual*, LCResources Inc., Lafayette, 1987

217 L.R. Snyder, J.W. Dolan & M.P. Rigney, *LC-GC*, 4 (1986) 921

218 L.R. Snyder, M.A. Quarry & J.L. Glajch, *Chromatographia*, 24 (1987) 970

219 J.W. Dolan & L.R. Snyder, *LC-GC*, 5 (1987) 970

220 J.W. Dolan, L.R. Snyder & M.A. Quarry, *Chromatographia*, 24 (1987) 261

221 F. Gago, J. Alvarez-Builla & J. Elguero, *J. Liq.*

Chromatogr., 10 (1987) 1031

222 J.E. Haky & A.M. Young, *J. Liq. Chromatogr.*, 7 (1984) 675

223 S.R. Bakalyar, R. McIlwrick & E. Roggendorf, *J. Chromatogr.*, 142 (1977) 353

224 R. Kaliszan, K. Osmialowski, S.A. Tomellini, S.H. Hsu, S.D. Fazio & R.A. Hartwick, *Chromatographia*, 20 (1985) 705

225 R. Kaliszan, K. Osmialowski, S.A. Tomellini, S-H. Hsu, S.D. Fazio & R.A. Hartwick, *J. Chromatogr.*, 352 (1986) 141

226 P.J. Schoenmakers, H.A. Billiet & L. de Galan, *J. Chromatogr.*, 218 (1981) 261

227 N. Tanaka, H. Goodell & B.L. Karger, *J. Chromatogr.*, 158 (1978) 233

APPENDIX 1

PRELIMINARY STUDIES ON THE USE OF RETENTION INDICES IN TETRAHYDROFURAN (THF)

Prior to the main study on retention indices a preliminary study was carried out on the use of retention indices in tetrahydrofuran (THF) containing eluents. Previous work at Loughborough had examined the effects of changing the methanol and acetonitrile concentration on the retention indices of a selection of compounds[1-3]. Although studies had been carried out using eluents containing selected THF proportions[4] or using THF as one component of a ternary eluent[5] no comparative study had been carried out using a range of eluent compositions. The aims were therefore to confirm the linearity of the log capacity factor - carbon number relationship for the alkylarylketones, to study the robustness of retention index values and to study the selectivity as compared to methanol and acetonitrile, THF being the third corner of the selectivity triangle frequently used in optimisation.

The experimental conditions were basically the same as those in described for the main study except that the aqueous phase was distilled water rather than a buffer. The column used was packed with a different batch (Spherisorb ODS batch 19/35) and loading (Spherisorb ODS rather than Spherisorb ODS2) to that used for the main study. The results obtained could therefore not be directly included in the retention prediction database.

Smith[2] has proposed a set of column test compounds which could be used to characterise column selectivity.

These compounds, methyl benzoate, p-cresol, 2-phenylethanol, nitrobenzene and toluene have been used in previous work on retention indices and were studied on the Spherisorb ODS column over the eluent ranges THF-H_2O 20:80 to 60:40.

The longest retained test compound, toluene, eluted before valerophenone so that a limited set of alkylarylketone standards (acetophenone to valerophenone) was used to calibrate the system.

a) Linearity of log k' - C_{no} relationship for alkylarylketones

The capacity factors of the alkylarylketone standards (Table 1) were used to calibrate the retention index scale using the linear relationship between log k' and 100 x C_{no}. At all the eluent compositions examined there is a linear relationship between log k' and carbon number (Table 2). The relationship between slope and THF concentration and intercept and THF concentration is not linear but can be described by the use of quadratic equations (Table 3), this was also found for acetonitrile in this study and has been reported for THF by Jandera[8].

b) Retention Indices of Test Compounds

The capacity factors (Table 1) were used to calculate the retention indices (Table 4) using the regression equations described above. The change in retention indices across the eluent range (Figure 1) was smaller than the change in capacity factors, however, all the compounds showed some change in retention index. The greatest changes

with eluent composition were observed with 2-phenylethanol and *p*-cresol, which are considerably more polar than the alkylarylketones. Toluene, which is less polar than the alkylarylketones, showed a steady increase in retention index with increasing THF concentration. As had been observed in acetonitrile and methanol containing eluents the retention index of methyl benzoate was almost unchanged over the eluent range. The retention indices obtained differed from those obtained in methanol and acetonitrile eluents[1,3,5] confirming the selectivity differences obtained using the different organic modifiers.

From these results it appears that the alkylarylketone scale is also suitable for use in THF containing eluents and the prediction scheme could also be expanded to include THF containing eluents.

Table 1: Capacity factors of alkylarylketone standards and column test compounds

Compound	Capacity factor Tetrahydrofuran proportion (%)				
	20	30	40	50	60
acetophenone	8.36	4.06	1.85	1.36	0.84
propiophenone	21.38	8.63	3.10	2.11	1.12
butyrophenone	50.82	16.33	4.62	2.93	1.38
valerophenone	128.97	31.16	6.81	4.00	1.66
p-cresol	19.19	7.81	2.51	1.70	0.87
methyl benzoate[a]	20.07	7.74	---	1.70	1.07
nitrobenzene	23.74	10.19	3.24	2.14	1.10
2-phenylethanol	6.21	2.95	1.31	0.97	0.60
toluene	65.69	22.54	5.99	3.84	1.77

a determined in a separate series of separations

Table 2: Relationship between log k' and 100 x carbon number in eluents containing tetrahydrofuran

$$\log k' = a\,(100 \times C_{no}) + b$$

Tetrahydrofuran proportion (%)	Coefficients of regression equations		correlation coefficient
	a	b	
20	3.943	-2.23	0.9999
30	2.931	-1.72	0.9992
40	1.867	-1.21	0.9973
50	1.540	-1.08	0.9966
60	0.977	-0.85	0.9948

Table 3: Relationship between slope and eluent composition and intercept and eluent composition in tetrahydrofuran containing eluents

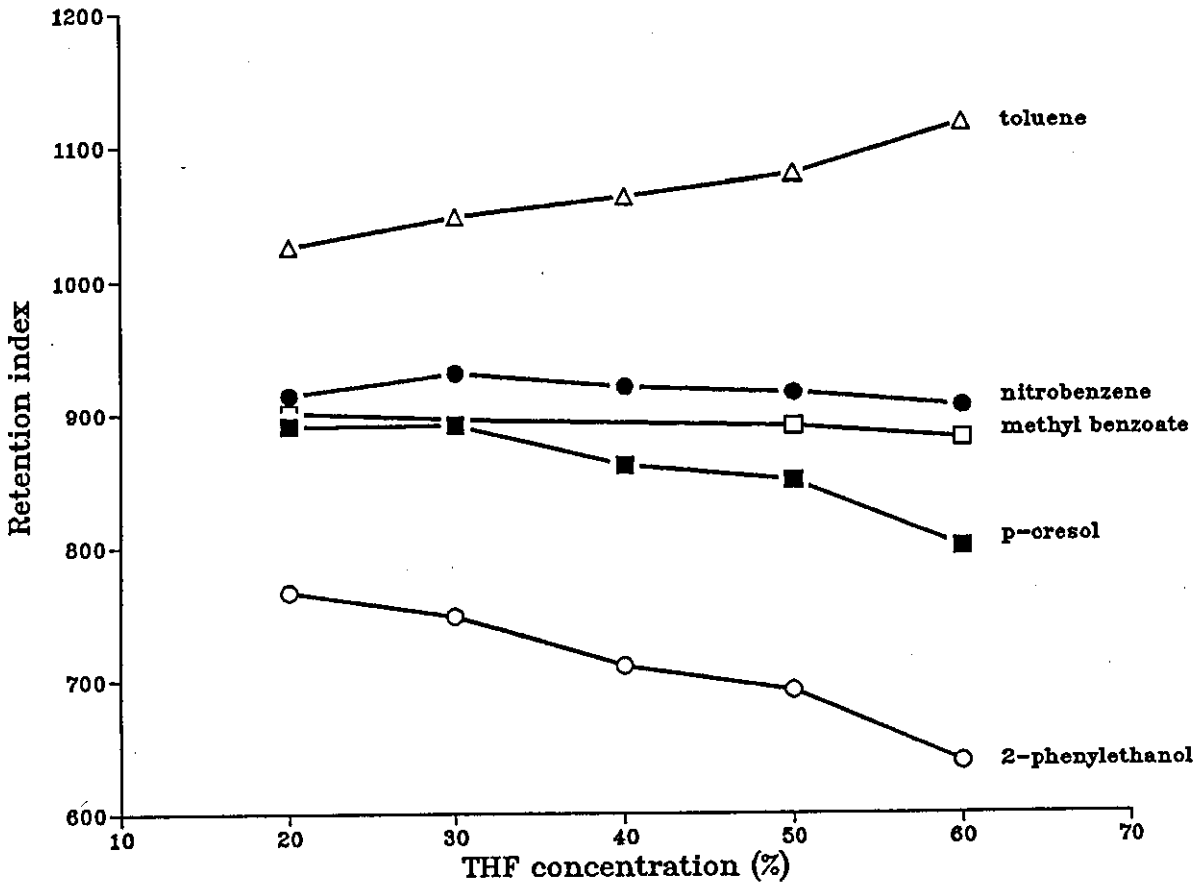
$$\text{slope} = ax^2 + bx + c \text{ or}$$
$$\text{intercept} = ax^2 + bx + c$$

	Coefficients of regression equations		
	a ($\times 10^6$)	b ($\times 10^3$)	c
Slope	1.168	-0.167	0.007
Intercept	-671.429	87.714	-3.718

Table 4: Retention indices of test compounds in acetonitrile eluents

Compound	Retention index Tetrahydrofuran (%)				
	20	30	40	50	60
p-cresol	891	892	862	851	801
methyl benzoate	901	896	---	892	883
nitrobenzene	914	931	921	917	907
2-phenylethanol	766	748	711	693	639
toluene	1026	1049	1064	1081	1119

Figure 1: Change in retention index of test compounds with tetrahydrofuran proportion



REFERENCES

1 R.M. Smith, G.A. Murilla and C.M. Burr *J. Chromatogr.* 388 (1987) 37

2 R.M. Smith, T.G. Hurdley, R. Gill and A.C. Moffat *Chromatographia* 19 (1984) 401

3 R.M. Smith *J. Chromatogr.* 236 (1982) 313

4 R.M. Smith *Anal. Chem.* 56 (1984) 256

5 R.M. Smith *J. Chromatogr.* 324 (1985) 243

6 P. Jandera *Chromatographia* 19 (1984) 101

APPENDIX 2

PUBLISHED WORK

R.M. Smith, G.A. Murilla and C.M. Burr, *J. Chromatogr.*, 388 (1987) 37
"Alkyl aryl ketones as a retention index scale with acetonitrile or tetrahydrofuran containing eluents in reversed-phase high performance liquid chromatography".

C.M. Burr and R.M. Smith, *Anal. Proc.* 25 (1988) 46
"Substituent interaction indices for retention prediction in reversed-phase high performance liquid chromatography".

C.M. Burr and R.M. Smith, *Anal. Proc.* submitted
"Retention prediction in RP-HPLC using a functional group database and expert system (CRIPES)".

R.M. Smith and C.M. Burr, *J. Chromatogr.* submitted
"Retention prediction of analytes in reversed-phase high preformance liquid chromatography based on molecular structure. Part 1. Monosubstituted aromatic compounds".

R.M. Smith and C.M. Burr, *J. Chromatogr.* submitted'
"Retention prediction of analytes in reversed-phase high preformance liquid chromatography based on molecular structure. Part II. Long term reproducibility of capacity factors and retention indices".