# COMPUTER-ASSISTED RETENTION PREDICTION SYSTEM FOR RE-VERSED-PHASE MICRO HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY 

KIYOKATSU JINNO* and KAZUYA KAWASAKI<br>School of Materials Science, Toyohashi University of Technology, Toyohashi 440 (Japan)


#### Abstract

SUMMARY

A retention prediction system for reversed-phase liquid chromatography was investigated. The system is based on the use of a retention description which was derived from studies of quantitative structure-retention relationships. A system for an octadecylsilica column was constructed on a 16 -bit microcomputer and then evaluated by comparing measured and predicted retention data. Excellent agreement between both values were observed. This system can also work well for any other reversed-phase columns, such as octyl-, phenyl- and ethylsilicas.


## INTRODUCTION

During the past decade, high-performance liquid chromatography (HPLC), especially reversed-phase HPLC (RPLC), has developed rapidly and is now accepted as a reliable and versatile analytical tool for the separation and quantitation of mixtures soluble in liquids. There is now a wide variety of commercially available instruments that completely satisfy the requirements of most routine analytical and semi-preparative separations. However, in practical separations, finding a good combination of the mobile phase and stationary phase is usually the most difficult and time-consuming task that faces chromatographers. This has traditionally been carried out by trial-and-error, based on personal experience and intuition, and therefore, could be facilitated if solute retention could be predicted under appropriate experimental conditions.

In order to predict the retention of any solute, a clear understanding of the retention mechanism will be needed. During recent years, much effort has been directed to investigating the mechanism of solute retention in RPLC, chromatographically and/or by spectroscopic measurements; major progress has been made ${ }^{1-4}$. At present, the "solvophobic theory" introduced by Horváth and co-workers ${ }^{5-8}$ is generally acknowledged as one of the most consistent theories available to describe solute distribution phenomena in RPLC. According to this entropically-driven interaction model, it can be anticipated that physicochemical parameters such as the solute's surface area, its partition coefficient between two immiscible phases and its aqueous solubility may correlate with the retention in RPLC. In practice, such correlations
exist and some of these parameters have been determined on the basis of those relationships. For example, many attempts have been made to use RPLC ${ }^{9-15}$ as the most promising method for the determination of the logarithm of the partition coefficient of a compound in the 1-octanol-water system, which is a good measure of its hydrophobicity; this is an important parameter in quantitative structure-activity relationships (QSARs).

The same idea as in QSAR can be applied to predict retention in RPLC. That is, the retention can be predicted on the premise that relationships exist between physicochemical parameters representing molecular properties, such as structure and/or shape, etc., and retention, if such parameters are available.

This paper is concerned with the construction of a retention prediction system (RPS) in RPLC for alkylbenzenes, polycyclic aromatic hydrocarbons (PAHs) and benzene derivatives substituted with polar functional groups on the basis of such established relationships.

## EXPERIMENTAL

The liquid chromatographic system consisted of a microfeeder MF-2 (Azuma Electric Co., Tokyo, Japan) as the pump and a Uvidec-100 III ultraviolet spectrophotometer (Jasco, Tokyo, Japan) as the detector, set at 210 or 260 nm . The column temperature was maintained at $20 \pm 0.1^{\circ} \mathrm{C}$ by a DW-620 thermostat (Komatsu, Tokyo, Japan).

Four reversed-phase columns (PTFE tubing, $12 \mathrm{~cm} \times 0.5 \mathrm{~mm}$ I.D.) were packed by the slurry technique; the packing materials were (a) Jasco FineSIL $\mathrm{C}_{18}$ ( 5 $\mu \mathrm{m}$ ), (b) Develosil (Nomura Chemicals, Seto, Japan) C $\mathrm{C}_{8}(5 \mu \mathrm{~m})$, (c) Develosil phenyl (C-p) $(10 \mu \mathrm{~m})$ and (d) FineSIL $\mathrm{C}_{2}(10 \mu \mathrm{~m})$.

The mobile phases comprised HPLC-grade methanol and acetonitrile (Kanto Chemicals, Tokyo, Japan), and purified water. All of the test substances were commercially available products from many sources, used as received. The flow-rate of the mobile phases was always $4 \mu \mathrm{l} / \mathrm{min}$.

The capacity factor, $k^{\prime}$, was evaluated from the retention time of the eluite, $t_{R}$, according to the equation $k^{\prime}=\left(t_{R}-t_{0}\right) / t_{0}\left(t_{0}=\right.$ retention time of an unretained compound), where a 50 ppm sodium nitrite solution was employed ${ }^{16,17}$ for $t_{0}$. All measurements were made in at least triplicate. The average reproducibility of each determination was better than $1.0 \%$ relative.

Multiple regression analyses were performed by the use of a MELCOM 800 computer (Mitsubishi Electric Co., Osaka, Japan). The retention prediction system (RPS) was constructed on a 16-bit microcomputer NEC PC-9800 (Nippon Electric Co., Tokyo, Japan) with a program written in BASIC language.

## RESULTS AND DISCUSSION

## Molecular structure descriptors

In order to establish the relationships between retention and parameters (here we say "descriptors") describing the molecular properties of a solute, it is necessary that those parameters have additive and constitutive properties which can be obtained not from experimental measurements but from theoretical calculations. This
is because such limitations can facilitate retention prediction with the aid of computer calculations.

The descriptors which satisfy the above requirements were selected from many published sources and those used in this study are as follows:
$\log P$, the logarithm of the partition coefficient in 1-octanol-water, which is a measure of the hydrophobicity of the molecule ${ }^{18,19}$
$\pi$, the hydrophobic substituent constant which is a measure of the hydrophobicity of the substituent group ${ }^{19}$

Van der Waals volume, $V_{\mathrm{w}}$, and Van der Waals surface area, $A_{\mathrm{w}}$, calculated from the Van der Waals radii of the atoms of which the molecule is composed ${ }^{20}$
$L / B$, the ratio of the maximized length to breadth of the rectangle enclosing the molecule and which indicates the shape of the molecule ${ }^{21,22}$

Molecular connectivity index, $\chi$, which shows the topological size of the molecule and its degree of branching ${ }^{23}$
$\sigma$, Hammett's constant
$H_{\mathrm{A}}, H_{\mathrm{D}}$, parameters showing the extent of substituent groups; $H_{\mathrm{A}}$ is equal to the number of hydrogen acceptor groups and $H_{\mathrm{D}}$ to the number of hydrogen donor groups in the molecule ${ }^{19}$

Correlation factor, $F$, calculated as (number of double bonds) + (number of primary and secondary carbon atoms) -0.5 for a non-aromatic ring ${ }^{24,25}$.

## Quantitative structure-retention relationships ${ }^{26}$

In the present study, the $k^{\prime}$ values of 20 alkylbenzenes, 18 PAHs and 28 benzene derivatives substituted with polar groups were determined on the four reversed-phase columns with mobile phases of methanol-water and acetonitrile-water.

The values of the descriptors and measured capacity factors of alkylbenzenes are listed in Table I. The correlation between $\log k^{\prime}$ and $F, \chi, V_{\mathrm{w}}, A_{\mathrm{w}}$ and $\log P$ was investigated by linear regression analysis. The correlation coefficients of each data set are summarized in Table II. These results can be expressed as follows:
(1) Almost all descriptors, except the correlation factor, $F$, show fairly good correlations with $\log k^{\prime}$, obtained in the eight systems, because $F$ was originally defined for PAHs and not for alkylbenzenes
(2) $\log P$, as shown in Table II, has the highest correlation coefficient with log $k^{\prime}$ among the five descriptors examined
(3) The greater the alkyl chain length of the bonded phase or the higher the carbon number of the bonded ligand, the higher is the correlation coefficient between $\log k^{\prime}$ and $V_{w}$ or $A_{w}$
(4) No differences between mobile phase systems were found.

These conclusions give a clear overview of the retention mechanism in RPLC. The hydrophobicity of a solute, which is measured by $\log P$, is the dominant factor controlling the retention, consistent with solvophobic theory. Further, in systems with the longer alkyl chain bonded phases, $V_{\mathrm{w}}$ and $A_{\mathrm{w}}$ of a solute contribute more to its retention than in systems with the shorter alkyl chain packings. These results lead to the following explanation. A solute can be held on the stationary phase with the longer alkyl chain by its hydrophobicity, and once it is held on the hydrophobic stationary phase, more energy might be required for a larger solute to escape from the hydrophobic interaction between it and that chain, because of the steric hindrance
TABLE I
CAPACITY FACTORS AND DESCRIPTORS OF ALKYLBENZENES ON FOUR REVERSED-PHASE COLUMNS

| Compound | Capacity factor, $k^{\prime}$, at $20^{\circ} \mathrm{C}$ |  |  |  |  |  |  |  | Descriptor |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Acetonitrile-water |  |  |  | Methanol-water |  |  |  | $F$ | $\chi$ | $V_{w}$ | $A_{w}$ | $\log P$ |
|  | $\begin{aligned} & 65: 35 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 50: 50 \\ & C-p \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{2} \end{aligned}$ | $\begin{aligned} & 75: 25 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 60: 40 \\ & C-p \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{2} \end{aligned}$ |  |  |  |  |  |
| Benzene | 0.98 | 1.82 | 1.61 | 1.50 | 0.81 | 1.50 | 1.02 | 0.88 | 3.0 | 2.000 | 48.36 | 6.01 | 2.13 |
| Toluene | 1.24 | 2.62 | 2.06 | 1.99 | 1.10 | 2.47 | 1.40 | 1.29 | 4.0 | 2.411 | 59.51 | 7.45 | 2.60 |
| Ethylbenzene | 1.53 | 3.80 | 2.65 | 2.69 | 1.41 | 3.97 | 1.92 | 1.99 | 5.0 | 2.971 | 69.74 | 8.80 | 3.13 |
| $o$-Xylene | 1.62 | 3.57 | 2.47 | 2.53 | 1.48 | 3.86 | 1.84 | 1.85 | 5.0 | 2.827 | 70.66 | 8.89 | 3.12 |
| $m$-Xylene | 1.60 | 3.88 | 2.57 | 2.60 | 1.56 | 4.20 | 1.88 | 1.99 | 5.0 | 2.821 | 70.66 | 8.89 | 3.12 |
| $n$-Propylbenzene | 2.02 | 5.78 | 3.39 | 3.75 | 1.95 | 6.74 | 2.79 | 3.15 | 6.0 | 3.471 | 79.97 | 10.15 | 3.66 |
| Isopropylbenzene | 1.88 | 5.32 | 3.33 | 3.48 | 1.80 | 6.09 | 2.58 | 2.83 | 5.0 | 3.354 | 79.96 | 10.14 | 3.54 |
| $o$-Ethyltoluene | 1.91 | 5.10 | 3.18 | 3.39 | 1.90 | 6.04 | 2.51 | 2.69 | 6.0 | 3.388 | 80.89 | 10.24 | 3.65 |
| $p$-Ethyltoluene | 2.03 | 5.59 | 3.34 | 3.67 | 2.08 | 6.72 | 2.68 | 3.13 | 6.0 | 3.382 | 80.89 | 10.24 | 3.65 |
| 1,2,4-Trimethylbenzene | 2.01 | 5.27 | 3.14 | 3.37 | 2.23 | 6.49 | 2.55 | 2.90 | 6.0 | 3.238 | 81.81 | 10.33 | 3.54 |
| $n$-Butylbenzene | 2.69 | 8.76 | 4.55 | 5.15 | 2.76 | 11.5 | 4.00 | 5.09 | 7.0 | 3.971 | 90.20 | 11.50 | 4.19 |
| sec.-Butylbenzene | 2.45 | 7.90 | 4.25 | 4.77 | 2.42 | 10.1 | 3.52 | 4.40 | 6.0 | 3.892 | 90.19 | 11.49 | 4.07 |
| $p$-Cymene | 2.44 | 7.74 | 4.11 | 4.80 | 2.60 | 10.4 | 3.57 | 4.62 | 6.0 | 3.765 | 91.11 | 11.58 | 4.06 |
| $m$-Diethylbenzene | 2.48 | 7.92 | 4.20 | 5.03 | 2.60 | 10.2 | 3.61 | 4.41 | 7.0 | 3.943 | 91.12 | 11.59 | 4.18 |
| 1,2,4,5-Tetramethylbenzene | 2.58 | 7.22 | 3.82 | 4.37 | 3.04 | 10.1 | 3.42 | 4.36 | 7.0 | 3.655 | 92.96 | 11.77 | 4.01 |
| $n$-Pentylbenzene | 3.60 | 13.2 | 5.97 | 7.14 | 3.96 | 19.6 | 5.96 | 8.16 | 8.0 | 4.471 | 100.43 | 12.85 | 4.72 |
| sec.-Pentylbenzene | 3.23 | 11.9 | 5.51 | 6.51 | 3.38 | 17.0 | 5.03 | 6.75 | 7.0 | 4.392 | 100.42 | 12.84 | 4.60 |
| Pentamethylbenzene | 3.10 | 9.09 | 4.44 | 5.32 | 4.14 | 14.3 | 4.49 | 5.90 | 8.0 | 4.077 | 104.11 | 13.21 | 4.48 |
| Hexamethylbenzene | 3.80 | 11.6 | 5.21 | 6.39 | 5.64 | 20.4 | 5.75 | 8.42 | 9.0 | 4.500 | 115.26 | 14.65 | 4.95 |
| $p$-Diisopropylbenzene | 3.67 | 15.2 | 6.46 | 8.39 | 4.31 | 25.2 | 6.82 | 10.1 | 7.0 | 4.708 | 111.56 | 14.27 | 5.00 |

TABLE II
RELATIONSHIPS BETWEEN $\log k^{\prime}$ AND DESCRIPTORS FOR ALKYLBENZENES

| Descriptor | Acetonitrile-water |  |  |  | Methanol-water |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{aligned} & 65: 35 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 50: 50 \\ & C-p \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{2} \end{aligned}$ | $\begin{aligned} & 75: 25 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 60: 40 \\ & C-p \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{2} \end{aligned}$ |
| $F$ | 0.948 | 0.904 | 0.889 | 0.897 | 0.968 | 0.925 | 0.920 | 0.992 |
| $\chi$ | 0.987 | 0.996 | 0.995 | 0.996 | 0.958 | 0.994 | 0.996 | 0.994 |
| $V_{w}$ | 0.987 | 0.969 | 0.959 | 0.967 | 0.992 | 0.985 | 0.980 | 0.983 |
| $A_{w}$ | 0.989 | 0.973 | 0.963 | 0.970 | 0.991 | 0.987 | 0.983 | 0.985 |
| $\log P$ | 0.995 | 0.992 | 0.987 | 0.991 | 0.979 | 0.997 | 0.996 | 0.996 |

due to the geometry of the longer alkyl chain. This is to say, the steric effect is well reflected in the solute retention. For the shorter alkyl chains, the steric hindrance is smaller than that observed for the longer ones, and hydrophobic interaction with the chain is the dominant factor governing solute retention ${ }^{27}$.

The correlation between $k^{\prime}$ values, measured on various phases, also show the existence of the above retention mechanism, a higher degree of correlation implying a similarity of the interaction in the systems for a given solute type. The results of linear regression analysis of each data set are tabulated as correlation matrices in Table III. High correlations are found between three columns, $\mathrm{C}_{2}, \mathrm{C}-\mathrm{p}$ and $\mathrm{C}_{8}$, while low correlations are found between $\mathrm{C}_{2}$ and $\mathrm{C}_{18}$, as well as C -p and $\mathrm{C}_{18}$. It is apparent that the interaction between the solutes and the $\mathrm{C}_{2}$ or C -p stationary phase is not entirely similar to the interaction between the solutes and the $\mathrm{C}_{18}$ stationary phase.

In view of the concept discussed above for alkylbenzenes, the description of retention by the use of the hydrophobic parameter, $\log P$, is a reasonable compromise.

The $k^{\prime}$ values of eighteen PAHs on the four reversed-phase columns were determined. The compositions of the mobile phases, methanol-water and acetonitrile-water, were chosen to give reasonable retention values. The results are shown in Table IV, which also lists the values of $F, \chi, V_{\mathrm{w}}, A_{\mathrm{w}}, \log P$ and $L / B$ for cach solute. The correlations between $\log k^{\prime}$ and these descriptors were investigated.

TABLE III
CORRELATION MATRICES FOR $\log k^{\prime}$ OF ALKYLBENZENES ON REVERSED-PHASE COLUMNS

|  | Acetonitrile-water |  |  |  | Methanol water |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{aligned} & 65: 35 \\ & \mathrm{C}_{18} \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{B} \end{aligned}$ | $\begin{aligned} & 50: 50 \\ & C-p \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{2} \end{aligned}$ | $\begin{aligned} & 75: 25 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 60: 40 \\ & c-p \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{2} \end{aligned}$ |
| $\mathrm{C}_{18}$ | 1.000 | 0.989 | 0.982 | 0.985 | 1.000 | 0.976 | 0.971 | 0.975 |
| $\mathrm{C}_{8}$ |  | 1.000 | 0.999 | 0.999 |  | 1.000 | 0.998 | 0.999 |
| C-p |  |  | 1.000 | 0.998 |  |  | 1.000 | 0.999 |
| $\mathrm{C}_{2}$ |  |  |  | 1.000 |  |  |  | 1.000 |

TABLE IV
CAPACITY FACTORS AND DESCRIPTORS OF PAHS ON FOUR REVERSED-PHASE COLUMNS

| Compound | Capacity factors, $k^{\prime}$, at $20^{\circ} \mathrm{C}$ |  |  |  |  |  |  |  | Descriptor |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Acetonitrile-water |  |  |  | Methanol-water |  |  |  | $F$ | $\chi$ | $\nu_{w}$ | $A_{w}$ | $\log P$ | $L / B$ |
|  | $\begin{aligned} & 65: 35 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 50: 50 \\ & C-p \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{2} \end{aligned}$ | $\begin{aligned} & 75: 25 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 60: 40 \\ & C-p \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{2} \end{aligned}$ |  |  |  |  |  |  |
| Indene | 1.36 | 2.97 | 2.40 | 2.41 | 1.34 | 2.94 | 1.90 | 1.49 | 4.5 | 3.211 | 68.07 | 7.77 | 3.11 | 1.22 |
| Indan | 1.75 | 4.29 | 2.92 | 3.13 | 1.78 | 4.81 | 2.41 | 2.20 | 5.5 | 3.534 | 72.41 | 8.47 | 3.46 | 1.16 |
| Naphthalene | 1.56 | 3.46 | 2.80 | 2.63 | 1.54 | 3.41 | 2.15 | 1.58 | 5.0 | 3.405 | 73.96 | 8.42 | 3.29 | 1.24 |
| Acenaphthylene | 1.87 | 4.15 | 3.37 | 3.40 | 1.95 | 4.38 | 2.72 | 1.83 | 5.5 | 4.149 | 83.44 | 8.84 | 3.74 | 1.08 |
| Acenaphthene | 2.25 | 5.44 | 3.84 | 3.78 | 2.74 | 6.84 | 3.78 | 2.83 | 6.5 | 4.445 | 87.88 | 9.54 | 4.09 | 1.06 |
| Anthracene | 2.96 | 6.62 | 4.85 | 4.43 | 3.75 | 8.39 | 5.01 | 3.41 | 7.0 | 4.809 | 99.56 | 10.84 | 4.45 | 1.57 |
| Phenanthrene | 2.74 | 6.15 | 4.58 | 4.09 | 3.35 | 7.60 | 4.45 | 3.02 | 7.0 | 4.815 | 99.56 | 10.84 | 4.45 | 1.46 |
| Pyrene | 4.15 | 8.58 | 5.71 | 5.03 | 5.65 | 12.00 | 6.42 | 4.17 | 8.0 | 5.559 | 109.04 | 11.26 | 4.90 | 1.27 |
| Fluoranthene | 3.74 | 8.09 | 5.65 | 5.18 | 5.04 | 11.00 | 6.07 | 4.19 | 8.0 | 5.565 | 109.04 | 11.26 | 4.90 | 1.22 |
| Benzo[b]fluorene | 4.55 | 10.40 | 7.20 | 6.81 | 7.65 | 16.90 | 9.64 | 7.40 | 8.5 | 6.017 | 119.27 | 12.61 | 5.43 | 1.78 |
| Naphthacene | 6.66 | 13.10 | 8.73 | 8.79 | 11.60 | 21.40 | 12.60 | 9.37 | 9.0 | 6.214 | 125.16 | 13.26 | 5.61 | 1.89 |
| Chrysene | 5.51 | 11.20 | 7.80 | 6.68 | 8.81 | 17.30 | 10.20 | 6.87 | 9.0 | 6.226 | 125.16 | 13.26 | 5.61 | 1.72 |
| Triphenylene | 5.04 | 10.30 | 7.30 | 6.04 | 7.91 | 15.80 | 8.86 | 5.59 | 9.0 | 6.232 | 125.16 | 13.26 | 5.61 | 1.12 |
| Benz[a]anthracene | 5.51 | 11.80 | 8.09 | 7.89 | 9.05 | 18.50 | 10.70 | 7.33 | 9.0 | 6.220 | 125.16 | 13.26 | 5.61 | 1.58 |
| Perylene | 8.33 | 14.60 | 9.48 | 7.69 | 11.70 | 25.40 | 14.00 | 8.73 | 10.0 | 6.976 | 134.64 | 13.68 | 6.06 | 1.27 |
| Benzo[a]pyrene | 9.21 | 15.90 | 9.74 | 8.69 | 16.90 | 28.60 | 15.10 | 10.20 | 10.0 | 6.970 | 134.64 | 13.68 | 6.06 | 1.50 |
| Benzo[ $]$ fluoranthene | 7.60 | 14.60 | 9.45 | 8.26 | 13.40 | 25.00 | 13.70 | 9.80 | 10.0 | 6.976 | 134.64 | 13.68 | 6.06 | 1.39 |
| Benzo[ $k]$ fluoranthene | 8.38 | 15.90 | 10.30 | 10.30 | 15.40 | 28.40 | 15.90 | 11.80 | 10.0 | 6.970 | 134.64 | 13.68 | 6.06 | 1.48 |

TABLE V
CORRELATION COEFFICIENTS FOR RELATIONSHIPS BETWEEN $\log k^{\prime}$ AND DESCRIPTORS FOR PAHs

| Descriptor | Acetonitrile-water |  |  |  | Methanol-water |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{aligned} & 65: 35 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 50: 50 \\ & C-p \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{2} \end{aligned}$ | $\begin{aligned} & 75: 25 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 60: 40 \\ & C-p \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{2} \end{aligned}$ |
| $F$ | 0.992 | 0.995 | 0.993 | 0.973 | 0.991 | 0.994 | 0.992 | 0.979 |
| $\chi$ | 0.990 | 0.990 | 0.993 | 0.973 | 0.989 | 0.989 | 0.991 | 0.974 |
| $V_{w}$ | 0.988 | 0.989 | 0.996 | 0.974 | 0.989 | 0.987 | 0.992 | 0.973 |
| $A_{\text {w }}$ | 0.978 | 0.984 | 0.992 | 0.971 | 0.983 | 0.982 | 0.988 | 0.973 |
| $\log P$ | 0.989 | 0.993 | 0.996 | 0.978 | 0.991 | 0.993 | 0.995 | 0.980 |
| $L / B$ | 0.513 | 0.538 | 0.555 | 0.590 | 0.550 | 0.537 | 0.567 | 0.596 |

Table V lists the correlation coefficients obtained by linear regression analysis. The results in Table V can be evaluated as follows:
(1) Almost all descriptors except $L / B$, particularly $F$ and $\log P$, show high correlations with $\log k^{\prime}$, measured in all systems.
(2) $L / B$ always shows poor correlations.
(3) The shorter the alkyl length of the bonded phase or the lower the carbon number of the bonded ligand, the higher is the correlation coefficient between $\log k^{\prime}$ and $L / B$.
(4) No differences were observed between the mobile phase systems.

As expected, the results (except item 4) indicate that the hydrophobicity and the size of a solute are the dominant factors controlling the retention in RPLC.

It is possible to say from these facts that the shape of a solute seems to be the dominant factor for PAH retention in RPLC. This means that the value of $L / B$ should be highly correlated with the retention. However, regression analysis shows the opposite. On the other hand, this observation in itself does not indicate that the shape of the solute is not important in controlling the retention.

Close examination of the data indicated that, although a higher correlation exists between $\log k^{\prime}$ vs. $F, \chi, V_{\mathrm{w}}, A_{\mathrm{w}}$ and $\log P$, these parameters could not explain the elution order of the structural isomers, such as naphthacene, chrysene and triphenylene, each of which has the same $F, V_{w}, A_{w}$ and $\log P$ values. However, the parameter $L / B$ can explain the elution order of these isomers on all reversed-phase columns. As shown in Table $V$, the contribution of $L / B$ to the solute retention shows an interesting trend, namely, as the alkyl chain length decreases the ability to distinguish those structural isomers increases. This tendency can be explained as follows. Both solutes with a rod-like shape, such as naphthacene, and solutes with a disc-like shape, such as triphenylene, can interact with the longer alkyl chain adequately. On the other hand, for the shorter alkyl chain the hydrophobic interaction with the solute having a disc-like shape is less complete than the interaction with the rod-shaped solute. This difference in the interaction is the primary reason for steric recognition with the shorter alkyl chain phases. Fig. 1 shows such a situation schematically; the $\mathrm{C}_{18}$ and $\mathrm{C}_{2}$ stationary phases are selected as examples for the longer and the shorter chain bonded phases, respectively.

Table VI shows correlation matrices for $\log k^{\prime}$, measured with various phases.


Fig. 1. Solute retention of naphthacene (II) and triphenylene (I) on $\mathrm{C}_{18}(\mathrm{~A})$ and $\mathrm{C}_{2}$ stationary phases (B). The partially shaded part in $B$ does not interact with the alkyl chains.

Higher correlation implies a similarity of the interaction in the systems for a given solute type, i.e., a similarity of retention mechanisms between column materials. Higher correlations are found between columns $\mathrm{C}_{8}$ and $\mathrm{C}_{18}, \mathrm{C}$-p and $\mathrm{C}_{18}$ and C -p and $\mathrm{C}_{8}$, while the lowest correlation exists between columns $\mathrm{C}_{2}$ and $\mathrm{C}_{18}$. These observations imply that the retention mechanisms of columns C-p and $\mathrm{C}_{8}$ may be considered to be similar, although some differences seem to exist between columns $\mathrm{C}_{18}$ and $\mathrm{C}_{2}$. This conclusion is in accord with the explanation given above.

From this discussion it is clear that the molecular size, shape and hydrophobicity of PAHs contribute to the control of their retention in RPLC. Therefore, stepwise multiple regression analyses were carried out for each data set in order to improve the correlation coefficients between $\log k^{\prime}$ and those descriptors. The results of this treatment are shown in Table VII. It is apparent from a comparison of Tables V and VII that multiple regression analysis is effective in improving the correlation coefficients for stationary phases with shorter chains, but it is not so effective for those with longer chains. However, it is clear that this multi-parameter description of the retention behaviour of PAHs is effective in all the systems examined.

TABLE VI
CORRELATION MATRICES FOR $\log k^{\prime}$ OF PAHs ON REVERSED-PHASE COLUMNS

|  | Acetonitrile water |  |  |  | Methanol water |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{aligned} & 65: 35 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 50: 50 \\ & C-p \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{2} \end{aligned}$ | $\begin{aligned} & 75: 25 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 60: 40 \\ & C-p \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{2} \end{aligned}$ |
| $\mathrm{C}_{18}$ | 1.000 | 0.995 | 0.992 | 0.979 | 1.000 | 0.996 | 0.997 | 0.991 |
| $\mathrm{C}_{8}$ |  | 1.000 | 0.996 | 0.988 |  | 1.000 | 0.997 | 0.993 |
| C-p |  |  | 1.000 | 0.989 |  |  | 1.000 | 0.993 |
| $\mathrm{C}_{2}$ |  |  |  | 1.000 |  |  |  | 1.000 |

TABLE VII
RELATIONSHIPS BETWEEN $\log k^{\prime}$ OF PAHs AND THE DESCRIPTORS $F$ AND $L / B$

| Column | Mobile phase | Equation | $n$ | $r$ | $s$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{C}_{18}$ | A | $\log k^{\prime}=0.142 F+0.050 L / B-0.596$ | 18 | 0.993 | 0.034 |
| $\mathrm{C}_{18}$ | M | $\log k^{\prime}=0.185 F+0.138 L / B-0.915$ | 18 | 0.994 | 0.041 |
| $\mathrm{C}_{8}$ | A | $\log k^{\prime}=0.123 F+0.073 L / B-0.155$ | 18 | 0.997 | 0.019 |
| $\mathrm{C}_{8}$ | M | $\log k^{\prime}=0.168 F+0.099 L / B-0.402$ | 18 | 0.996 | 0.029 |
| C-p | A | $\log k^{\prime}=0.106 F+0.084 L / B-0.191$ | 18 | 0.997 | 0.017 |
| C-p | M | $\log k^{\prime}=0.158 F+0.147 L / B-0.633$ | 18 | 0.997 | 0.026 |
| $\mathrm{C}_{2}$ | A | $\log k^{\prime}=0.094 F+0.123 L / B-0.745$ | 18 | 0.983 | 0.038 |
| $\mathrm{C}_{2}$ | M | $\log k^{\prime}=0.145 F+0.193 L / B-0.846$ | 18 | 0.989 | 0.046 |

TABLE VIII
RETENTION DESCRIPTION BY COMBINATIONS OF DESCRIPTORS FOR ALKYLBENZENES AND PAHs

| Column | Mobile phase | Equation | $n$ | $r$ | $s$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{C}_{18}$ | A | $\log k^{\prime}=0.187 \log P+0.029 F-0.552$ | 38 | 0.990 | 0.037 |
| $\mathrm{C}_{18}$ | M | $\log k^{\prime}=0.230 \log P+0.056 F-0.847$ | 38 | 0.989 | 0.054 |
| $\mathrm{C}_{8}$ | A | $\log k^{\prime}=0.070 A_{w}+0.067 \log P-0.331$ | 38 | 0.983 | 0.045 |
| $\mathrm{C}_{8}$ | M | $\log k^{\prime}=0.107 A_{\mathrm{w}}+0.095 \log P-0.634$ | 38 | 0.989 | 0.049 |
| C-p | A | $\log k^{\prime}=0.235 \log P-0.022 F-0.213$ | 38 | 0.995 | 0.022 |
| C-p | M | $\log k^{\prime}=0.325 \log P-0.016 F-0.663$ | 38 | 0.996 | 0.028 |
| $\mathrm{C}_{2}$ | A | $\log k^{\prime}=0.049 A_{w}+0.090 \log P-0.284$ | 38 | 0.979 | 0.042 |
| $\mathrm{C}_{2}$ | M | $\log k^{\prime}=0.010 A_{w}+0.056 \log P-0.829$ | 38 | 0.982 | 0.056 |

TABLE IX
CAPACITY FACTORS AND DESCRIPTORS OF SUBSTITUTED BENZENE DERIVATIVWES ON FOUR REVERSED-PHASE COLUMNS

| Compound | Capacity factor, $k^{\prime}$, at $20^{\circ} \mathrm{C}$ |  |  |  |  |  |  |  | Descriptor |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Acetonitrile-water |  |  |  | Methanol-water |  |  |  | $F$ | $\chi$ | $V_{w}$ | $A_{w}$ | $\pi$ | $H_{A}$ | $H_{D}$ |
|  | $\begin{aligned} & 65: 35 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 50: 50 \\ & C-p \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{2} \end{aligned}$ | $\begin{aligned} & 75: 25 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 60: 40 \\ & C-p \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{2} \end{aligned}$ |  |  |  |  |  |  |  |
| Aniline | 0.63 | 0.70 | 0.90 | 0.74 | 0.56 | 0.48 | 0.67 | 0.36 | 3.0 | 2.199 | 56.38 | 7.07 | -1.23 | 1 | I |
| N -Ethylaniline | 1.13 | 1.80 | 1.77 | 1.83 | 0.96 | 1.28 | 1.33 | 0.88 | 5.0 | 3.221 | 77.82 | 9.79 | 0.08 | , | 1 |
| Benzaldehyde | 0.68 | 1.03 | 1.12 | 1.08 | 0.60 | 0.74 | 0.88 | 0.48 | 3.0 | 2.435 | 60.06 | 7.61 | -0.65 | 1 | 0 |
| Benzonitrile | 0.73 | 1.17 | 1.30 | 1.27 | 0.59 | 0.73 | 0.95 | 0.56 | 3.0 | 2.384 | 60.54 | 7.52 | $-0.57$ | 1 | 0 |
| Nitrobenzene | 0.83 | 1.45 | 1.51 | 1.64 | 0.77 | 1.06 | 1.13 | 0.68 | 3.0 | 2.448 | 62.64 | 7.88 | -0.28 | 1 | 0 |
| Anisole | 0.91 | 1.71 | 1.56 | 1.75 | 0.88 | 1.37 | 1.20 | 0.88 | 4.0 | 2.523 | 62.71 | 7.99 | -0.02 | 1 | 0 |
| Acetophenone | 0.72 | 1.04 | 1.11 | 1.06 | 0.68 | 0.85 | 0.99 | 0.58 | 4.0 | 2.865 | 71.21 | 9.05 | -0.55 | 1 | 0 |
| Methyl benzoate | 0.91 | 1.50 | 1.43 | 1.55 | 0.86 | 1.38 | 1.34 | 0.91 | 4.0 | 2.977 | 76.73 | 9.83 | -0.01 | 1 | 0 |
| $\alpha$-Bromotoluene | 1.12 | 2.75 | 2.42 | 2.68 | 1.03 | 2.32 | 1.86 | 1.29 | 4.0 | 3.667 | 70.47 | 8.76 | 0.79 | 0 | 0 |
| $p$-Nitroaniline | 0.54 | 0.78 | 1.04 | 0.82 | 0.54 | 0.47 | 0.68 | 0.32 | 3.0 | 2.647 | 70.66 | 8.94 | -1.51 | 2 |  |
| $m$-Chloroaniline | 0.74 | 1.20 | 1.35 | 1.17 | 0.66 | 0.90 | 0.95 | 0.52 | 3.0 | 2.712 | 65.86 | 8.20 | -0.52 | 1 | 1 |
| $p$-Ethylaniline | 1.05 | 1.25 | 1.37 | 1.17 | 0.99 | 1.14 | 1.25 | 0.66 | 5.0 | 3.171 | 77.76 | 9.86 | -0.21 | 1 | 1 |
| $m$-Aminoacetophenone | 0.52 | 0.55 | 0.73 | 0.54 | 0.48 | 0.38 | 0.68 | 0.27 | 4.0 | 3.064 | 79.23 | 10.11 | -1.78 | 2 | 1 |
| Diethyl phthalate | 1.02 | 1.99 | 1.96 | 2.03 | 0.87 | 1.60 | 1.92 | 1.17 | 7.0 | 5.135 | 125.56 | 16.35 | 1.02 | 2 | 0 |
| $o$-Nitrotoluene | 1.01 | 1.93 | 1.86 | 1.79 | 0.92 | 1.57 | 1.48 | 0.98 | 4.0 | 2.865 | 73.79 | 9.32 | 0.28 | 1 | 0 |
| $\alpha$-Bromo-p-nitrotoluene | 1.05 | 2.27 | 2.40 | 2.44 | 0.99 | 1.73 | 1.87 | 1.08 | 4.0 | 4.115 | 84.75 | 10.63 | 0.51 | 1 | 0 |

TABLE X
CORRELATION COEFFICIENTS FOR RELATIONSHIPS BETWEEN $\log k^{\prime}$ AND DESCRIPTORS FOR SUBSTITUTED BENZENE DERIVATIVES

| Descriptor | Acetonitrile-water |  |  |  | Methanol-water |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{aligned} & 65: 35 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 50: 50 \\ & C-p \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{2} \end{aligned}$ | $\begin{aligned} & 75: 25 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 60: 40 \\ & C-p \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{2} \end{aligned}$ |
| $F$ | 0.571 | 0.429 | 0.419 | 0.389 | 0.538 | 0.480 | 0.622 | 0.535 |
| $\chi$ | 0.536 | 0.546 | 0.593 | 0.526 | 0.517 | 0.553 | 0.731 | 0.588 |
| $V_{w}$ | 0.376 | 0.344 | 0.378 | 0.322 | 0.351 | 0.365 | 0.561 | 0.423 |
| $A_{\text {w }}$ | 0.360 | 0.325 | 0.356 | 0.304 | 0.336 | 0.350 | 0.545 | 0.410 |
| $\pi$ | 0.924 | 0.961 | 0.941 | 0.949 | 0.891 | 0.967 | 0.976 | 0.982 |
| $H_{\text {A }}$ | -0.516 | -0.523 | -0.471 | -0.516 | -0.503 | -0.543 | $-0.374$ | -0.484 |
| $H_{\text {D }}$ | -0.374 | -0.574 | -0.510 | -0.611 | $-0.363$ | -0.569 | -0.553 | -0.612 |
| $H_{\mathrm{A}}+H_{\mathrm{D}}$ | -0.544 | -0.672 | -0.600 | -0.690 | -0.530 | -0.681 | -0.568 | -0.672 |
| $H_{\mathrm{A}}-H_{\mathrm{D}}$ | -0.123 | -0.044 | 0.033 | 0.083 | -0.121 | 0.023 | 0.155 | 0.111 |

It was found that the hydrophobicity, size and shape of the solutes were the dominant factors controlling the retention of PAHs, as well as of alkylbenzenes. This is because both alkylbenzenes and PAHs are similar in hydrocarbon nature, i.e., they do not have polar functional groups. Therefore, in order to introduce a more general retention description which holds for both PAHs and alkylbenzenes, stepwise multiple regression analyses were performed. At the same time, the correlations between descriptors were evaluated, and the results indicated that they were very highly correlated with each other. Thus, in statistical terms, the combination of these descriptors was not preferable and did not affect partial regression coefficients variance. However, such a combination is able to improve the accuracy of retention prediction. The results of these analyses are in Table VIII. For the $\mathrm{C}_{18}$ and C-p packings, retention descriptions by combination of $\log P$ and $F$ give higher correlation coefficients than those by either parameter, alone. For the $\mathrm{C}_{8}$ and $\mathrm{C}_{2}$ packings, the descriptions with $\log P$ and $A_{\mathrm{w}}$ are the most effective.

To our knowledge ${ }^{28-30}$, it is common in this kind of study that substituted benzenes should be divided into two groups, such as phenols and others. We therefore first examined non-phenolic compounds. The correlation analyses between measured capacity factors and several descriptors were carried out in the same manner as for alkylbenzenes and PAHs. The required data are summarized in Table IX. In this casc, $\pi$ was used as a measure of the hydrophobicity of the compounds, because benzene is the parent compound of each of them.

Table X lists the correlation coefficients between $\log k^{\prime}$ and each descriptor, obtained by linear regression analysis. It is clear that only the correlation between $\log k^{\prime}$ and $\pi$ is fairly high, but that other descriptors show poor correlations. This means that the size and shape of the solute is not dominant in controlling its retention and that the hydrophobicity is the only dominant factor. It will be noted that different results were observed in the case of alkylbenzenes and PAHs, where the size and shape of molecules contributed significantly to their retention.

Stepwise multiple regression analysis allowed the introduction of a second descriptor, ( $\left.H_{\mathrm{A}}-H_{\mathrm{D}}\right)$. This descriptor indicates the hydrogen-accepting ability of the
TABLE XI
RELATIONSHIPS BETWEEN $\log k^{\prime}$ OF SUBSTITUTED BENZENE DERIVATIVES AND THE DESCRIPTORS $\pi$ AND ( $H_{\mathrm{A}}-H_{\mathrm{D}}$ )

| Column | Mobile phase | Equation | $n$ | $r$ | $s$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{C}_{18}$ | A | $\log k^{\prime}=0.136 \pi-0.049\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)-0.007$ | 16 | 0.959 | 0.034 |
| $\mathrm{C}_{18}$ | M | $\log k^{\prime}=0.130 \pi-0.048\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)-0.050$ | 16 | 0.925 | 0.045 |
| $\mathrm{C}_{8}$ | A | $\log k^{\prime}=0.240 \pi-0.031\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)+0.215$ | 16 | 0.965 | 0.054 |
| $\mathrm{C}_{8}$ | M | $\log k^{\prime}=0.283 \pi-0.045\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)+0.117$ | 16 | 0.978 | 0.043 |
| C-p | A | $\log k^{\prime}=0.178 \pi-0.025\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)+0.222$ | 16 | 0.976 | 0.035 |
| C-p | M | $\log k^{\prime}=0.191 \pi+0.005\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)+0.106$ | 16 | 0.982 | 0.031 |
| $\mathrm{C}_{2}$ | A | $\log k^{\prime}=0.234 \pi-0.017\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)+0.212$ | 16 | 0.950 | 0.064 |
| $\mathrm{C}_{2}$ | M | $\log k^{\prime}=0.273 \pi-0.010\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)-0.098$ | 16 | 0.982 | 0.041 |

whole molecule. Table XI shows the retention descriptions for substituted benzene derivatives. It is apparent that the regression coefficients of ( $H_{\mathrm{A}}-H_{\mathrm{D}}$ ) in the equations are always negative for all systems except the C -p-methanol system. This fact can be explained as follows. Molecules for which $\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)$ is positive have a high hydrogen-accepting ability, and show a tendency to prefer the protonic hydrogenrich phase, i.e., the polar mobile phase. Thus, the retention of such molecules should be weaker than that of any solute for which $\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)$ is not positive, even if it has the same hydrophobicity as the former. However, these descriptions are not effective for the $\mathrm{C}-\mathrm{p}$ and $\mathrm{C}_{2}$ stationary phases. In those cases, fairly good correlations were obtained by the use of $\pi$ only. The reasons are, first, the effect of the presence of a large number of unreacted silanols is more serious for stationary phases with shorter alkyl chains than for phases with longer alkyl chains, due to the difference in the steric hindrance of the ligands, which was observed in the experiments as tailing chromatographic peaks on the $\mathrm{C}_{2}$ and $\mathrm{C}-\mathrm{p}$ phases. This is because the penetration of the solute into the phase down to the surface is easier with short alkyl chains. Secondly, the solvation of the stationary phase with the polar mobile phase with the aid of the hydrogen-bonding ability of residual silanols occurs more easily with the shorter alkyl chains than with the longer alkyl chains. Therefore, for stationary phases with longer chains, the protonic hydrogen-rich phase is only the mobile phase, and the contribution of $\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)$ to the solute retention is fairly high, while for the shorter chain phases the stationary phase partially changes to be more like the mobile phase, because of the solvation, and the hydrophobic interaction between the solute and the ligands is only the dominant factor in the retention. Then, the contribution of ( $H_{\mathrm{A}}-H_{\mathrm{D}}$ ) to the retention is small or negligible. Those facts are also clearly indicated in the correlation matrices shown in Table XII. The lack of similarity between $\mathrm{C}_{18}$ and $\mathrm{C}_{2}$ and $\mathrm{C}-\mathrm{p}$ in retention mechanism is reflected in the low correlation between these phases.

Our correlation study on phenols indicated that no single descriptor could describe the retention of these compounds, and that it required the use of Hammett's parameter as the second descriptor ${ }^{30}$. The measured retention data and descriptors for phenols are summarized in Table XIII.

The same procedure as that used in the case of non-phenolic compounds was used for phenols, and the results of the linear regression analysis for each data set are in Table XIV. For both mobile phase systems it is apparent that the correlations between columns $\mathrm{C}_{2}$ and $\mathrm{C}_{8}$, and $\mathrm{C}-\mathrm{p}$ and $\mathrm{C}_{18}$, are high, while those between columns

TABLE XII
CORRELATION MATRICES FOR $\log k^{\prime}$ OF SUBSTITUTED BENZENE DERIVATIVES ON RE-VERSED-PHASE COLUMNS

|  | Acetonitrile-water |  |  |  | Methanol-water |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{aligned} & 65: 35 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 50: 50 \\ & C-p \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{2} \end{aligned}$ | $\begin{aligned} & 75: 25 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 60: 40 \\ & C-p \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{2} \end{aligned}$ |
| $\mathrm{C}_{18}$ | 1.000 | 0.922 | 0.899 | 0.896 | 1.000 | 0.945 | 0.923 | 0.913 |
| $\mathrm{C}_{8}$ |  | 1.000 | 0.986 | 0.993 |  | 1.000 | 0.964 | 0.986 |
| C-p |  |  | 1.000 | 0.971 |  |  | 1.000 | 0.982 |
| $\mathrm{C}_{2}$ |  |  |  | 1.000 |  |  |  | 1.000 |

TABLE XIII
CAPACITY FACTORS AND DESCRIPTORS OF PHENOLS ON FOUR REVERSED-PHASE COLUMNS

| Compound | Capacity factor, $k^{\prime}$, at $20^{\circ} \mathrm{C}$ |  |  |  |  |  |  |  | Descriptor |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Acetonitrile-water |  |  |  | Methanol-water |  |  |  | $\pi$ | $\sigma$ |
|  | $\begin{aligned} & 65: 35 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 50: 50 \\ & C-p \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{2} \end{aligned}$ | $\begin{aligned} & 75: 25 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 60: 40 \\ & C-p \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{2} \end{aligned}$ |  |  |
| Phenol | 0.44 | 0.72 | 0.84 | 0.54 | 0.45 | 0.62 | 0.65 | 0.43 | 0 | 0 |
| $o$-Cresol | 0.59 | 1.00 | 1.08 | 0.93 | 0.57 | 0.99 | 0.87 | 0.63 | 0.51 | -0.10 |
| $m$-Aminophenol | 0.31 | 0.34 | 0.52 | 0.33 | 0.33 | 0.23 | 0.48 | 0.24 | -1.29 | -0.16 |
| $p$-Aminophenol | 0.70 | 0.34 | 0.78 | 0.38 | 0.75 | 0.30 | 0.91 | 0.32 | -1.63 | -0.66 |
| $o$-Nitrophenol | 0.73 | 1.25 | 1.34 | 1.22 | 0.72 | 0.94 | 1.00 | 0.62 | 0.33 | 2.00 |
| $m$-Nitrophenol | 0.53 | 0.84 | 1.06 | 0.84 | 0.57 | 0.84 | 0.82 | 0.61 | 0.54 | 0.71 |
| $p$-Nitrophenol | 0.49 | 0.77 | 0.98 | 0.88 | 0.55 | 0.74 | 0.81 | 0.54 | 0.50 | 0.78 |
| $m$-Ethylphenol | 0.64 | 1.26 | 1.26 | 1.12 | 0.64 | 1.37 | 1.16 | 0.85 | 0.94 | -0.07 |
| $o$-Chlorophenol | 0.57 | 0.98 | 1.09 | 0.99 | 0.57 | 0.95 | 0.90 | 0.63 | 0.69 | 0.40 |
| $p$-Chlorophenol | 0.61 | 1.09 | 1.18 | 0.97 | 0.68 | 1.24 | 1.02 | 0.71 | 0.93 | 0.23 |
| $p$-Methoxyphenol | 0.42 | 0.59 | 0.75 | 0.61 | 0.43 | 0.50 | 0.67 | 0.41 | -0.12 | -0.16 |
| $p$-Hydroxyacetanilide | 0.27 | 0.27 | 0.42 | 0.25 | 0.32 | 0.23 | 0.42 | 0.14 | -0.97 | 0 |

## TABLE XIV

CORRELATION MATRICES FOR $\log k^{\prime}$ OF PHENOLS ON REVERSED-PHASE COLUMNS

|  | Acetonitrile water |  |  |  | Methanol-water |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{aligned} & 65: 35 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 50: 50 \\ & C-p \end{aligned}$ | $\begin{aligned} & 55: 45 \\ & C_{2} \end{aligned}$ | $\begin{aligned} & 75: 25 \\ & C_{18} \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{8} \end{aligned}$ | $\begin{aligned} & 60: 40 \\ & C-p \end{aligned}$ | $\begin{aligned} & 65: 35 \\ & C_{2} \end{aligned}$ |
| $\mathrm{C}_{18}$ | 1.000 | 0.701 | 0.887 | 0.751 | 1.000 | 0.650 | 0.946 | 0.749 |
| $\mathrm{C}_{8}$ |  | 1.000 | 0.944 | 0.978 |  | 1.000 | 0.822 | 0.962 |
| C-p |  |  | 1.000 | 0.961 |  |  | 1.000 | 0.897 |
| $\mathrm{C}_{2}$ |  |  |  | 1.000 |  |  |  | 1.000 |

$\mathrm{C}_{2}$ and $\mathrm{C}_{18}$, and $\mathrm{C}_{8}$ and $\mathrm{C}_{18}$, are low. These results are almost consistent with the facts observed in our other investigations, that is, some dissimilarity exists in the retention mechanism between the shorter and the longer chain phases. The higher correlation between columns C-p and $\mathrm{C}_{18}$ for phenols than that for substituted benzene derivatives may be due to the fact that the $\pi \pi$ interaction between the phenyl groups of phenols and those of the C-p phase is stronger than that between substituted benzenes and the C-p phase. The observed dissimilarity in the retention mechanism is mainly due to the difference in the effect of the steric hindrance of the corresponding ligands.

Multi-combination treatments were performed for phenols on various stationary phases. The results are shown in Table XV. In these retention descriptions the physical meaning of $\pi \sigma(1-\pi)$ is not clear yet. However, close examination of the data in Table XIII implies that a molecule which has a high possibility to form intra-molecular hydrogen bonds shows larger retention than that of other molecules which have the same substituent groups but cannot form intra-molecular hydrogen bonds. A good example is the case of nitrophenols. o-Nitrophenol forms intra-molecular hydrogen bonds, but $m$ - and $p$-nitrophenols do not. These different types of behaviour of the compounds can be distinguished by Hammett's constant. In this instance, $\sigma$ of $o$-nitrophenol is 2.00 and the $\sigma$ values of $m$ - and $p$-nitrophenols are 0.71 and 0.78 , respectively.

Finally the quantitative structure-retention relationships discussed here gave highly correlated relationships between the retention data of alkylbenzenes, PAHs and substituted benzenes and several physicochemical descriptors, and these established relationships can be applied to construct a retention prediction system for those compound groups.

## System construction

As there is a highly correlated relationship between $\log k^{\prime}$ and several descriptors for aromatic compounds, as discussed in the previous section, the following $N$ equations can be obtained by multiple regression analyses for $N$ experimental conditions

$$
\begin{array}{cc}
X=X_{1} & \log k_{1}^{\prime}=g_{1} D_{1}+h_{1} D_{2}+i_{1} \\
X=X_{2} & \log k_{2}^{\prime}=g_{2} D_{1}+h_{2} D_{2}+i_{2} \\
X=X_{3} & \log k_{3}^{\prime}=g_{3} D_{1}+h_{3} D_{2}+i_{3} \\
\vdots & \vdots  \tag{1}\\
X=X_{N} & \log k_{N}^{\prime}=g_{N} D_{1}+h_{N} D_{2}+i_{N}
\end{array}
$$



| 2LOO | 9560 | 21 |  | W | ${ }^{\text { }}$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 9 LO 0 | ES60 | 21 |  | V | ² |
| SEOO | 0260 | 21 |  | W | d-3 |
| $8+00^{\circ}$ | 0960 | 21 |  | V | d-J |
| 6500 | $6 L 60$ | 21 |  | W | ${ }^{8} 5$ |
| $8 \mathrm{LO} 0^{\circ}$ | 2560 | 21 |  | V | ${ }^{8}$ |
| OEOO | +L60 | 21 |  | W | ${ }^{15}$ |
| Sto 0 | 2S60 | 21 |  | V | ${ }^{81}$ |
| $s$ | $\downarrow$ | $u$ | uoupnb ${ }^{\text {a }}$ | asbyd ${ }^{\text {pll }}$ \%ow | ииumo |

where $X$ is the volume fraction of organic solvent in the mobile phase, $D_{1}$ and $D_{2}$ are descriptors of the retention, such as $\log P$ and $F$ for alkylbenzenes and PAHs, $\pi$ and $\left(H_{\mathrm{A}}-H_{\mathrm{D}}\right)$ for non-phenolic substituted benzenes and $\pi$ and $\pi \sigma(1-\pi)$ for phenols with the $\mathrm{C}_{18}$ column, respectively, and $g, h$ and $i$ are constants. If $g, h$ and $i$ can be expressed as a function of $X$, i.e., $X$ and $g, X$ and $h$ and $X$ and $i$ are well correlated, the following three equations can be obtained by multiple regression analyses

$$
\begin{align*}
& g=\mathrm{f}_{1}(X)=\sum_{m=0}^{n} j X^{m}  \tag{2}\\
& h=\mathrm{f}_{2}(X)=\sum_{m=0}^{n} k X^{m}  \tag{3}\\
& i=\mathrm{f}_{3}(X)=\sum_{m=0}^{n} l X^{m} \tag{4}
\end{align*}
$$

where $n$ is a number less than $N$, and $j, k$ and $l$ are constants. Multiple regression analyses have been carried out for selected standards with the $\mathrm{C}_{18}$ columns, and they gave the highly correlated relationships of $\mathrm{f}_{1}(X), \mathrm{f}_{2}(X)$ and $\mathrm{f}_{3}(X)$ for alkylbenzenes, PAHs and substituted benzenes.

The next type of equation for each data set is obtained as shown in Table XVI:

$$
\begin{equation*}
\log k^{\prime}=\mathrm{f}_{1}(X) D_{1}+\mathrm{f}_{2}(X) D_{2}+\mathrm{f}_{3}(X) \tag{5}
\end{equation*}
$$

Eqn. 5 means that, if $X$, the concentration of organic solvent in the mobile phase, and the descriptors $D_{1}$ and $D_{2}$, i.e., the characteristics of a solute, are known, the logarithm of the capacity factor, $\log k^{\prime}$, can be determined for given chromatographic conditions. This is the basic concept of RPS investigated here. As a typical example for constructing a RPS, we have selected the $\mathrm{C}_{18}$ column in the following discussions. Given eqn. 5 for the $\mathrm{C}_{18}$ column, it is possible to construct a RPS with the aid of a microcomputer for alkylbenzenes, PAHs and substituted benzene derivatives. The flow diagram of this scheme is shown in Figs. 2 and 3 for the computer-assisted RPS proposed here. The system was constructed on a 16 -bit microcomputer with programs for calculating several descriptors of alkylbenzenes, PAHs and substituted benzene derivatives.


Fig. 2. Basic scheme for the retention prediction system (RPS) in RPLC.


Fig. 3. Flow chart for retention prediction in RPLC.


















$\qquad$

Na : ©
No. : Coherabiz
Ne. 3 - $\mathrm{HHACH}_{4} \mathrm{CH} 3 \mathrm{OH} \mathrm{H}$



*+*** PLEASE SET IP THE SEPAFATIOM CONGITIDAS *****


Columb: Finesil cis.es


fow : ate: 4 al mari

theoretu:2l elate rumber: zot
ruEnlt output ir. Na * 。

| compound | k' | retention time (min) |
| :---: | :---: | :---: |
| $\mathrm{COH5C2H5}$ | 3.77 | 23.86 |
| COHJC (CH3)3 | 0.84 | 39.20 |
| $\mathrm{C} 6 \mathrm{H} 4 \mathrm{CH}(\mathrm{CH} 3) 2 \mathrm{CH}(\mathrm{CH} 3) 2$ | 13.37 | 71.87 |


| *** SEparimtion conoltions *** | SAMPLE LIST |
| :---: | :---: |
| columf : \#iriesil cle-s | 1. CEHESTHE |
|  | $\therefore$ Cohs ce $\mathrm{CH}_{3} 3$ |
| eluptit : H3CM:HEL = 50 : fot | 3. Coh4 $(\mathrm{H} / \mathrm{CH} 3) 2 \mathrm{CH}(\mathrm{CH} 3)=$ |
| 910e tate: 4 alimin |  |
| celd tame: 5 (mar |  |
| theoreticas plate number: ziou |  |

$; 1$


Fig. 4. Example of the function of RPS in predicting retention times of three alkylbenzenes.

| ************************ | RETENTION DATA | ************************* |
| :---: | :---: | :---: |
| compound | k' | retention time (min) |
| benzene | 1.04 | 10.19 |
| naphthalene | 2.07 | 15.36 |
| acenaphthylene | 2.63 | 18.16 |
| acenaphthene | 3.38 | 21.89 |
| fluorene | 3.64 | 23.20 |
| phenanthrene | 4.13 | 25.67 |
| pyrene | 5.53 | 32.66 |
| chrysene | 8.25 | 46.24 |
| perylene | 11.04 | 60.19 |
| dibenz (a,h)anthracene | 16.46 | 87.28 |
| *********************** | ************ | ************************ |



Fig. 5. Example of access for the retention prediction of ten PAHs.

## System access

Using the interactive mode, the following data should be input in the RPS: (i) the compound name or chemical formula of the desired solutes; (ii) the kind of mobile phase (mobile phase selection); (iii) the volume fraction of organic modifier in the eluent. The system also has the functions to give the optimum experimental conditions, such as analysis time, mobile phase composition, etc., for separating some particular solutes of interest ${ }^{31}$. Upon accessing the system, the RPS responds with the numerical capacity factor values and retention times for the desired solutes, the optimized separation conditions and a display of the idealized chromatogram. Typical examples are shown in Figs. 4 and 5. In Fig. 4 the chromatogram synthesized with the RPS is compared with the determined chromatogram. It is apparent that the performance of the RPS is excellent.

For other packing materials, the required procedure to make the RPS is to obtain some retention data for standard substances (four or five compounds) on every column one wants to use at various mobile phase compositions, e.g.,
methanol-water ranging from $80: 20$ to $30: 70$. The derived equations should be stored in the memory of the microcomputer. Then these equations can be used for retention prediction with the program at any time until the column deteriorates. The RPS has, of course, the functions to construct the system for each column from only the inputed retention data of the selected standards at appropriate mobile phase compositions.

## Prediction of unknowns

In order to illustrate the predictive ability of the RPS, it was used to predict the retention of a number of alkylbenzenes, PAHs and substituted benzene derivatives. Comparisons of predicted and measured $k^{\prime}$ values are shown in Figs. 6 and 7. It is apparent that the predictive ability of the system is quite satisfactory.

## Practical aspects

In identification procedures with chromatographic techniques it is common practice to compare the retention of standard materials with that of the peaks observed in the sample chromatogram. Thus, it is necessary to have some standard substances, even if they are highly toxic and add pollution problems. Also, chromatograms of practical samples are normally very complicated, because a great num-


Fig. 6. Relationship between predicted and observed $k^{\prime}$ values with a mobile phase of acetonitrile-water (65:35) on $\mathrm{C}_{18}$ stationary phase.


Fig. 7. Relationship between predicted and observed $k^{\prime}$ values with a mobile phase of methanol-water (75:25) on $\mathrm{C}_{18}$ stationary phase.
ber of species are present. This makes identification without high-resolution techniques difficult. These disadvantages can be obviated by using a RPS. The rapid identification of solutes without standard substances is non-polluting ${ }^{32,33}$. Focusing on some selected species, identified by the RPS, more precise information could be obtained by more advanced and complex analytical techniques, such as gas chromatography-mass spectrometry (GC-MS) with a spectra database system, but other spectroscopic techniques, such as Fourier transform-IR or UV (with a photodiode array for three-dimensional measurements), together with a RPS would have even greater potential in the identification of practical samples.

In the current RPS, only alkylbenzenes, PAHs and benzene derivatives substituted with polar functional groups can be accessed to predict retention with two mobile phase systems, aqueous acetonitrile and methanol. Expanding the applicability of this system to other compound groups and more complicated mobile phase systems, such as ternary eluents and/or gradient elution, will open new dimensions for liquid chromatography as a more powerful analytical technique.

## ACKNOWLEDGEMENTS

We express our sincere thanks to Drs. Y. Hirata and C. Fujimoto for their fruitful discussions through our work. This work was supported in part by The Asahi Glass Foundation For Industrial Technology (Tokyo, Japan).

## REFERENCES

1 C. R. Yonker, T. A. Zwier and M. F. Burke, J. Chromatogr., 241 (1982) 257.
2 H. Hemetsberger, P. Behrensmeyer, J. Henning and H. Ricken, Chromatographia, 12 (1979) 71.
3 L. C. Sander, J. B. Callis and L. R. Field, Anal. Chem., 55 (1983) 1068.
4 R. H. Gilpin and M. E. Gangoda, J. Chromatogr. Sci., 21 (1983) 352.
5 Cs. Horváth, W. Melander and I. Molnár, J. Chromatogr., 125 (1976) 129.
6 W. Melander, J. Stoveken and Cs. Horváth, J. Chromatogr., 199 (1980) 35.
7 Cs. Horváth and W. Melander, Amer. Lab., 10 (1978) 17.
8 Cs. Horváth (Editor), High-Performance Liquid Chromatography -Advances and Perspectives, Vol. 2, Academic Press, New York, 1980.
9 M. S. Mirrless, S. J. Moulton, C. T. Murphy and P. J. Taylor, J. Med. Chem., 19 (1976) 615.
10 H. Könemann, R. Zelle, F. Busser and W. E. Hammers, J. Chromatogr., 178 (1979) 559.
11 K. Miyake and H. Terada, J. Chromatogr., 240 (1982) 9.
12 K. Miyake and H. Terada, J. Chromatogr., 157 (1978) 386.
13 C. L. K. Lins, J. H. Block, R. J. Doerge and G. J. Barnes, J. Pharm. Sci., 71 (1982) 614.
14 G. L. Biagi, A. M. Barbaro, M. C. Guerra, G. Hakein, G. C. Solaini and P. A. Borea, J. Chromatogr., 177 (1979) 35.
15 K. Jinno, Chromatographia, 15 (1982) 667.
16 K. Jinno, N. Ozaki and T. Sato, Chromatographia, 17 (1983) 341.
17 K. Jinno, Chromatographia, 17 (1983) 367.
18 R. F. Rekker, The Hydrophobic Fragmental Constant, Elsevier, Amsterdam, 1977.
19 C. Hansch and A. Leo, Substituted Constants for Correlation Analysis in Chemistry and Biology, Wiley, New York, 1979.
20 A. Bondi, J. Phys. Chem., 68 (1964) 441.
21 S. A. Wise, W. J. Bonnett, F. R. Guenther and W. E. May, J. Chromatogr. Sci., 19 (1981) 457.
22 A. Radecki, H. Lamparczyk and R. Kaliszan, Chromatographia, 12 (1979) 595.
23 L. B. Kier and L. H. Hall, Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York, 1976.
24 J. F. Schabron, R. J. Hurtubise and H. F. Silver, Anal. Chem., 49 (1977) 2253.
25 R. J. Hurtubise, T. W. Allen and H. F. Silver, J. Chromatogr., 235 (1982) 517.
26 T. L. Hafkenscheid, Ph.D. Thesis, University of Amsterdam, Amsterdam, 1983.
27 K. Jinno and K. Kawasaki, Chromatographia, 17 (1983) 337.
28 K. Jinno and K. Kawasaki, Chromatographia, in press.
29 K. Jinno and K. Kawasaki, Chromatographia, 18 (1984) 103.
30 K. Jinno and K. Kawasaki, Chromatographia, 18 (1984) 90.
31 K. Jinno and K. Kawasaki, J. Chromatogr., 298 (1984) 326.
32 K. Jinno and K. Kawasaki, Chromatographia, 18 (1984) 211.
33 K. Jinno, Y. Hirata, H. Ohta, S. Sasaki and H. Abe, Anal. Lett., 17 (A10) (1984) 905.

