Calculation of retention in reversed-phase liquid chromatography

IV. ChromDream software for the selection of initial conditions and for simulating chromatographic behaviour

S.V. Galushko*, A.A. Kamenchuk and G.L. Pit

Institute of Bioorganic Chemistry and Oil Chemistry, Academy of Sciences of the Ukraine, 253660 Kiev 94 (Ukraine)

ABSTRACT

A computer-aided method to assist the chromatographer in the determination of initial conditions in reversed-phase HPLC has been developed. ChromDream software predicts the optimum methanol, acetonitrile or tetrahydrofuran concentration in a mobile phase on the basis of the structural formulae of the compounds. The effects of the concentration of an organic solvent in water on retention and chromatograms of compound mixtures are simulated. Reversed-phase column parameters responsible for selectivity and retention are calculated. Volume and energy increments for more than 100 volume fragments and more than 30 bond dipoles have been derived. An approach for the selection of reference compounds for column calibration is proposed.

INTRODUCTION

The selection of the initial conditions is the first step in HPLC method development and the possibility of using relationships between chemical structure and retention can save a lot of time and effort in this stage. Several software packages and expert systems for the calculation of retention in reversed-phase HPLC on the basis of chemical structure have been developed. An integrated expert system in HPLC contains LABEL and DASH expert systems for predictions of the initial conditions in reversed-phase HPLC [1]. LABEL suggests possible mobile phase conditions on a cyano-bonded column on entering the molecular size and polarity of the

analyte [2]. DASH predicts the percentage of methanol in a mobile phase for capacity factors of basic drugs between 3 and 10 to be obtained [3,4]. The user is required to translate the structure of an analyte into fragments. Each fragment is linked to a percentage of methanol and the final concentration of methanol is calculated by summation of the fragment contributions.

The CRIPES expert system for calculating retention indices based on the alkyl aryl ketone scale has been described [5]. The retention index was calculated by summation of the retention index of a parent compound and different contributions. The user was required to enter analyte details such as the type and the numbers of aliphatic and aromatic substituents.

METABOLEXPERT and ELUEX software packages calculate the capacity factor (k') of an

---

* Corresponding author.

analyte [6,7]. The programs are based on the
assumption that log k' in reversed-phase (C_{18})
HPLC is linearly related to the calculated
logarithm of the partition coefficient of an ana-
lyte in the octanol–water system (log P) and that
only small differences exist between reversed-
phase columns produced by different manufac-
turers. To calculate approximate log P values
fragmental constants were used.

Empirical correlation approaches are normally
not associated with model concepts of the re-
tention mechanism, the surface layer structure
and the character of interactions of retained
substances with mobile and stationary phases.
Previously we described a retention model and
an approach for the calculation of the retention
in reversed-phase HPLC (RP-HPLC) on the
basis of the structural formula of a compound
[8]. The possibility of predicting the retention of
compounds on different octadecyl columns with
water, water–methanol, water–acetonitrile and
water–tetrahydrofuran mobile phases has been
shown [8–12].

The aim of this work was to develop a flexible
computer-aided method for selecting the initial
conditions in reversed-phase HPLC and predict-
ing the chromatographic behaviour of com-
pounds on the basis of their structural formulae
and calibrating reversed-phase columns.

MODEL

To calculate approximately the retention, a
two-layer continuum model of a reversed-phase
chromatographic system was employed [8]:

(1) The surface of a modifier sorbent in RP-
HPLC has a surface layer (SL) that involves
octadecyl radicals and some of the components
of a mobile phase (MP).

(2) The SL is assumed to be a quasi-liquid
that has its own characteristics, such as surface
tension (γ_s) and dielectric constant (ε_s). The SL
characteristics vary with varying MP composition
and sorbent properties.

(3) The molecules of retained substances
penetrate into the SL. The retention is deter-
mined by the difference in molecule solvation
energies in the MP and SL.

To calculate the solvation energies, a molecule

of a substance is considered to consist of volume
fragments and bond dipoles, each of which
interact separately with the surrounding con-
tinuum.

Hence the model takes into account several
important points: (i) individuality of a packing;
(ii) the surface and polarity of a solute; and (iii)
the influence of mobile phase composition on the
interaction of a solute with the stationary and
mobile phases.

The model and approaches for calculating the
energy of solvation and sorbent parameters have
been described in detail and the equation de-
rived for calculating the retention in reversed-
phase HPLC and calibrating the column is
[8,10]:

$$\ln k'_x = a \left(\sum_i V_i \right)^{2/3} + b \left(\sum_j \Delta G_{\text{e.s.}j,\text{H}_2\text{O}} \right) + c \qquad (1)$$

where V_i are the increments of the partial molar
volumes of fragments in water, $G_{\text{e.s.}j,\text{H}_2\text{O}}$ are the
increments of energy of interaction of bond
dipoles with water, a, b and c are parameters of
the reversed-phase chromatographic system and
the column, including the phase ratio, and some
characteristics of the SP and MP [8,10]. Each
volume fragment i has its energy increment of
the $G_{\text{e.s.}i,\text{H}_2\text{O}}$. Hence the energy of interaction of
a molecule with water can be expressed in the
form

$$\sum_j \Delta G_{\text{e.s.}j,\text{H}_2\text{O}} = \sum_i \Delta G_{\text{e.s.}i,\text{H}_2\text{O}} + \sum_f \Delta G_{\text{e.s.}f,\text{H}_2\text{O}}$$

$$(2)$$

where $\Delta G_{\text{e.s.}f,\text{H}_2\text{O}}$ are increments of the contribu-
tion of $G_{\text{e.s.H}_2\text{O}}$ for dipoles that are not included
in volume fragments.

EXPERIMENTAL

Chromatographic conditions

The experiments were performed on an LKB
(Bromma, Sweden) liquid chromatographic sys-
tem consisting of a Model 2140 rapid spectral
detector set at 254 nm, a Model 2150 HPLC
pumping system, a Model 2152 LC controller, a
Model 2155 HPLC column oven, a Model 2154
injector and a Model 2220 recording integrator.

The columns used were (1) Polyol-SI 100, 5 μm (250 × 4.6 mm I.D.) (Serva, Heidelberg, Germany), (2) Nucleosil 120-5 C_{18} (100 × 4.0 mm I.D. and 30 × 4.0 mm I.D. cartridge) (Macherey–Nagel, Duren, Germany), (3) Eurospher 80-5 C_{18} (250 × 4.0 mm I.D.) (Eurochrom Knauer, Berlin, Germany), all stainless-steel columns, and (4) Separon SGX, 5 μm glass column (150 × 3.0 mm I.D.) (Tessek, Prague, Czech Republic). The mobile phase was phosphate buffer (pH 7.0; 0.01 M H_3PO_4 adjusted with 0.01 M sodium hydroxide) at a flow-rate of 0.5 ml/min. Separations were performed at 35°C for columns 1 and 4 and at 22°C for columns 2 and 3.

Software

Computer simulations were carried out using ChromDream software (Dr. Ing. H. Knauer GmbH, Berlin, Zehlendorf, Germany). The program was written in PASCAL language. The program was run on an IBM-compatible Piramida computer.

RESULTS AND DISCUSSIONS

Translating structure into fragments

To calculate the retention, a molecule must be translated into volume fragments and bond dipoles and subsequently increments V_i and also $\Delta G_{e.s.i,H_2O}$ and $\Delta G_{e.s.f,H_2O}$ must be used. One of the major difficulties is to provide a good user interface. It has been found earlier that when a user manually translates structures into fragments more than 50% of structures are incorrectly translated [4]. The conclusion was drawn that to avoid these mistakes a concise user manual is necessary in which it is explained how a structure can be translated into structure elements and additionally how the number of elements must be restricted [4]. Of course, the best alternative is to use a program to enter first the complete chemical structure and then to translate it automatically into fragments and to calculate the retention and simulate chromatographic behaviour and chromatograms. We have developed such a program and a structural formula drawn on the PC screen is translated automatically into fragments. This procedure enables us to employ many structural elements and to calculate the retention more precisely. Moreover, the program checks if a drawn structural formula fulfils the valency rules, indicates mistakes and determines some topological features of a compound: aromatic, cyclic, heterocyclic, saturated, etc. After drawing a structure on the screen, the software can analyse the structure immediately and display the list of fragments (see Table VIII) or save the structure in a current library file (Sample) for further simulating chromatograms or chromatographic behaviour. Up to 100 structures may be saved in each Sample and names of 100 Sample files may be displayed in the Sample Menu. A menu of alphanumeric and Boolean fields with topological features, different functional groups and heteroatoms allows a formula or a group of formulae to be found in a current library file.

Volume and energy parameters

Previously we derived volume and energy increments for some basic fragments and dipoles of organic compounds [8,10]. Possibilities of automatically translating a structural formula enabled us to expand the set of fragments and dipoles substantially. To obtain the values of V_i, experimental partial molar data for more than 300 compounds were used [14]. Experimental data have shown that the additivity of the action of separate molecule fragments is a good approximation for calculating partial molar volumes [8,14]. V_i values for the majority of the fragments were calculated from the experimental V values by the least-squares method and the rest were derived from Van der Waals volumes. The values of the increments obtained (Table I) enabled us to calculate the partial molar volumes of compounds belonging to the different classes with relative standard deviations of not more than 5% (Table II). It should be noted that the molecular volume of a compound is a popular parameter in liquid chromatography and we believe that it is useful to have the possibility for calculating it rapidly in the chromatographic software.

The energy of interaction of a molecule with water ($\Delta G_{e.s.,H_2O}$) is calculated as described previously [8]. As a first approximation, incre-

TABLE I

INCREMENTS OF PARTIAL MOLAR VOLUMES (V, IN cm^3 mol^{-1}) AND ENERGY ($\Delta G_{e.s.i,H_2O}$, IN kJ mol^{-1}) FOR SOME FRAGMENTS

Fragment[a]	V	$-\Delta G_{e.s.i,H_2O}$	Fragment[a]	V	$-\Delta G_{e.s.i,H_2O}$
H	5.00	0.00	C=O	16.00	32.20
-C-	-7.01	0.00	C=O(ar)	11.00	33.20
cyC	-7.00	0.00	cyC=O	11.00	27.50
=C=	0.00	0.00	arC=O	13.00	20.00
=C-	-3.69	0.00	COOH	24.00	65.00
ar=C-	5.16	0.00	COOH(ar)	22.84	62.63
cy=C-	5.00	0.00	CHO	20.73	35.00
C_{sp}	7.00	0.00	CHO(ar)	22.23	32.00
CH	3.50	1.42	CN	17.74	31.25
cyCH	4.59	1.42	CN(ar)	17.74	18.25
ar=CH-	11.87	4.36	CONH$_2$	31.38	71.68
=CH-	15.00	4.36	CONH$_2$(ar)	27.90	60.00
cy=CH-	9.00	4.36	F	12.00	0.00
CH$_{in}$	12.00	9.84	F(ar)	14.00	0.00
CH$_2$	15.68	2.84	CF$_3$	36.00	5.00
CH$_2$(ar)	16.68	-2.00	CF$_3$(ar)	32.00	2.00
cyCH$_2$	15.17	2.84	Cl	24.00	0.00
=CH$_2$	22.00	2.84	Cl(ar)	24.00	0.00
CH$_3$	27.43	4.26	Br	29.00	0.00
CH$_3$(ar)	23.43	2.26	Br(ar)	29.00	0.00
CH$_4$	34.00	5.68	I	36.00	0.00
-N-	-5.08	-8.00	I(ar)	36.00	0.00
-N-(ar)	-5.08	-3.00	-P-	-3.00	2.00
=N-	-2.00	0.00	=P-	0.00	2.00
ar=N-	5.89	15.00	cy=P-	8.00	2.00
arN	-4.00	0.00	cyP	-3.00	2.00
cyN	-2.50	0.00	P=O	20.00	48.00
cy=N-	6.00	0.00	cyP=O	21.00	48.00
NH	6.15	19.20	=PO-	26.00	48.00
NH(ar)	6.15	15.20	cy=PO-	27.00	48.00
cyNH	6.66	12.00	POOH	37.00	70.00
arNH	11.74	17.20	OH(ar)	12.20	33.00
NHp	3.15	15.00	cyPOOH	38.00	70.00
NH(ar)p	3.15	14.50	PO(OH)$_2$	44.00	95.00
cyNHp	6.15	9.50	S	7.14	-1.00
arNHp	6.15	13.00	S(ar)	7.14	-1.00
NH$_2$	13.00	44.00	Sarar	7.14	-1.00
NH$_2$(ar)	15.00	36.00	S(co)	7.14	-1.00
=NO-	14.00	38.00	=S	9.00	0.00
N=O	10.00	38.00	cyS	11.00	-1.00
cy=NO-	15.00	38.00	SH	15.00	19.00
cyN=O	15.00	38.00	S=O	21.43	42.00
NO$_2$	20.33	24.00	cyS=O	22.00	42.00
NO$_2$(ar)	20.33	20.00	SO$_2$	28.00	60.00
O	4.40	-4.00	cySO$_2$	29.00	60.00
O(co)	1.40	0.00	SO$_2$-NH$_2$	40.00	100.00
O(COar)	1.40	-14.00	SO$_2$NH$_2$ar	40.00	96.00
O(ar)	3.40	5.50	SO$_3$H	41.00	100.00
Oarar	3.40	-2.00	SO$_3$Har	41.00	100.00
cyO	6.00	-3.00	SOOH	34.00	42.00
OH	12.20	38.80	cyI	10.01	0.00

[a] cy = fragment in a ring; ar = fragment in an aromatic ring; (ar) = fragment connected to an aromatic ring; C_{sp} = fragment C≡; CH$_{in}$ = fragment HC≡; (co) = fragment connected to -CO- group; (COar) = fragment connected to CO(ar) ...

TABLE II

COMPARISON OF CALCULATED AND EXPERIMENTAL VALUES OF PARTIAL MOLAR VOLUME (V, IN $cm^3 mol^{-1}$) IN WATER

Compound	V(calc.)	V(exp.) [13]
1-Pentanol	102.35	102.60
2-Pentanol	101.92	102.60
3-Pentanol	101.92	101.20
Benzyl alcohol	103.40	102.20
Sorbitol	118.56	119.20
Phenol	86.72	86.20
Cyanophenol	97.75	98.00
Anisole	105.35	105.60
Dimethyloxolane	110.39	111.00
Cyclohexanedione	100.69	92.80
Cyclohexanone	100.86	99.70
Benzimidazole	97.31	98.53
1,10-Phenanthroline	137.39	142.42
Quinazoline	104.19	104.20
Benzylamine	104.20	104.70
Hexanic acid	114.10	116.00
2-Phenylacetic acid	115.20	114.10
Glutaric acid	95.00	99.14
Acetanilide	121.30	121.30
Ribose	98.34	95.20
Glucose	115.60	111.70

TABLE III

INCREMENTS OF $\Delta G_{e.s.j,H_2O}$ (IN kJ mol^{-1}) FOR SOME DIPOLES

Dipole	$-\Delta G_{e.s.j,H_2O}$	Dipole	$-\Delta G_{e.s.j,H_2O}$
C_{sp}–H	9.84	S–N	2.00
C_{sp^2}–H	4.36	P–H	5.00
C_{sp^3}–H	1.42	C_{sp}–Cl	2.00
C_{sp^2}–C_{sp^3}	2.33	C_{sp^2}–Cl	3.00
C_{sp^2}–C_{sp}	6.68	C_{sp^3}–Cl	23.00
C_{sp^3}–C_{sp}	11.06	C_{sp}–F	4.00
C_{sp}–O	1.15	C_{sp^2}–F	9.00
C_{sp^2}–O	2.15	C_{sp^3}–F	18.00
C_{sp^3}–O	12.00	C_{sp}–Br	1.00
C_{sp}–N	1.00	C_{sp^2}–Br	3.50
C_{sp^2}–N	2.00	C_{sp_3}–Br	22.00
C_{sp^3}–N	12.00	C_{sp}–I	1.00
C_{sp^2}=N	1.24	C_{sp^2}–I	2.50
O–N	0.71	C_{sp^3}–I	20.00
O–P	0.70	C_{sp^2}–S	1.00
O–S	0.50	C_{sp^3}–S	4.00
C_{sp^3}–P	2.00	C_{sp^2}=S	2.00
C_{sp^2}–P	4.00		

ments $\Delta G_{e.s.j,H_2O}$ for bond dipoles have been used [8]. By analysing the deviations in calculating the retention of more than 400 compounds belonging to different classes we found the values of $\Delta G_{e.s.j,H_2O}$ and $\Delta G_{e.s.i,H_2O}$ increments that allow more exact results (Tables I and III). Of course, the values of $\Delta G_{e.s.j,H_2O}$ increments can be further refined.

The program searches for volume fragments and dipoles in the structural formula and calculates $\Sigma_i (V_i)^{2/3}$ and $\Sigma_j \Delta G_{e.s.j,H_2O}$ according to eqns. 1 and 2. Sometimes a fragment can be in several dipoles (e.g., the fragment Cl can be involved in three dipoles: C_{sp^3}–Cl, C_{sp^2}–Cl and C_{sp}–Cl). In this case the program identifies the type of hybridization of a carbon atom bonded with a halogen atom and selects a corresponding value from the list of the parameters (Table III).

An opinion exists that "correct prediction . . . cannot be expected for a wide range of . . . compounds" [1]. On the other hand, experimental data show that fairly correct predictions can

be obtained for a certain family of substances, packing type and organic solvent in a mobile phase [4,5]. As will be shown further, a fairly large set of volume fragments and dipoles (Tables I and III) allows the satisfactory prediction of retention and acceptable mobile phase composition for various classes of substances [13]. However, it is possible that for a certain family of substances energy increments allowing more precise results can be obtained. To make the program more flexible, a special file PARAM.DAT is available. This file contains volume and energy parameters (Tables I and III) in a usual numerical format and these parameters may be changed by the user. Moreover, new dipoles can be added. For instance one can add an ar=C–Cl dipole to the list of the dipoles in the parameters file (Table III). In this case the program will differentiate a C_{sp^2}–Cl dipole from ar=C–Cl. Any fragment from Table I can be used to create new dipoles. Thus, in addition to the general system of the parameters derived (Tables I and III), one can develop several other more specific systems for certain classes or families of substances. The program allows the user to calculate the retention and to simulate a

chromatogram of a mixture of compounds using different systems of parameters simultaneously.

Column calibration

Eqn. 1 gives an opportunity to calibrate columns and to compare different reversed-phase columns and packings. By using several reference compounds, parameters a, b and c responsible for retention and selectivity can be calculated. In the framework of the approach used, the parameters a and b are determined by the differences in surface tension and dielectric permittivity of a sorbent surface layer and a mobile phase [8,10]. Hence these parameters enable us to evaluate column-to-column reproducibility and differences between columns quantitatively. The smaller the differences in parameters a, b and c, the smaller are the differences in the retention and selectivity of the separation. It has been shown that the parameters depend on bonded-phase structure, coverage density and concentration of silanol groups on the surface of a packing material [15]. The more silanol groups in the bonded layer, the higher are its surface tension and dielectric constant [15].

As each packing and column has a unique set of parameters [8,10,11,15], some procedure should be performed to determine the parameters. To calculate the column parameters (eqn. 1), capacity factor values for several reference compounds are entered and the software calculates the parameters a, b and c by the least-squares method. After determination of the column parameters, the capacity factors of the reference compounds are calculated to check the

validity of the parameters and to evaluate deviations between experimental and calculated values of the capacity factors (Table IV). Of course, the proposed method, as with any semi-empirical method, is an approximation and some deviations should be expected in results of column calibrations and retention calculations when different sets of reference compounds are used. Hence to compare different columns more objectively, the same eluent and reference compounds should be used. One of the suitable sets of reference compounds for column calibration consists of phenol, benzene, *p*-hydroxybenzaldehyde and toluene (or chlorobenzene) (Table IV). Other sets may be chosen in order to obtain capacity factors of several compounds in either one or several runs.

One of the problems in HPLC is the limited stability of chromatographic columns. To evaluate reversed-phase HPLC columns, 50–60% methanol–water mobile phases and several test mixtures have been proposed [16,17]. We consider that the calculation of column parameters with the ChromDream software using retention data for these or other mixtures can be a rapid, sensitive and quantitative test procedure both for the evaluation of new reversed-phase columns and in daily routine work for column stability control.

Parameters for different columns (Hypersil ODS, Nucleosil 120-5 C_{18}, Spherisorb ODS-2, Eurospher 80-5 C_{18}, ODS Si-100 Polyol, Separon SGX C_{18}, etc.) have been calculated for a wide range of methanol, acetonitrile and tetrahydrofuran concentrations in the mobile phase. These

TABLE IV

RESULTS OF CALIBRATING A COLUMN (NUCLEOSIL 120-5 C_{18}, 10 cm × 4 mm I.D.)

Mobile phase: methanol–water (50:50). $a = 0.2756$; $b = 0.0562$; $c = -2.2020$. V in $cm^3\ mol^{-1}$; $-\Delta G_{e.s.H_2O}$ in kJ mol^{-1}; k'_e = experiment capacity factor values; k'_c = calculated capacity factor values.

Compound	V	$-\Delta G_{e.s.H_2O}$	Ln k'_e	Ln k'_c	k'_e	k'_c
Chlorobenzene	98.52	24.80	2.27	2.28	9.69	9.82
Benzene	81.23	26.16	1.47	1.50	4.37	4.50
Toluene	97.95	26.39	2.20	2.17	9.00	8.78
Phenol	86.22	56.95	0.03	0.00	1.03	1.00
p-Hydroxybenzaldehyde	101.74	85.59	−0.94	−0.93	0.39	0.40

parameters are employed by the program for simulating chromatograms and calculating retention under selected conditions (column, concentration of an organic solvent in water, etc.). In addition to these parameters, the user can determine parameters for other reversed-phase columns and eluents and save them in addition to a description of chromatographic conditions in a parameters library for further chromatogram simulation and retention calculation.

Selecting reference compounds

As has been described above, to calibrate a column the capacity factor values for several reference compounds must be entered. In many instances parameters that were determined using compounds belonging to one class of substances are acceptable for calculating the retentions of substances belonging to other classes (see Fig. 4) [14]. However, to obtain more correct predictions for a certain class or family of substances, several compounds belonging to the family can be used for column calibration. To determine column parameters more precisely, it is necessary to employ compounds that have substantial differences in $V^{2/3}$ and $\Delta G_{e.s.}$ (eqn. 1), *i.e.*, that differ in surface and polarity. As the number of reference compounds rises, the error in determining column parameters decreases. Hence the selection of an adequate set of reference compounds is an important point in a column calibration. To select reference compounds that represent a family of substances adequately, we propose to use a surface in $V^{2/3}$–$\Delta G_{e.s.}$ coordinates. In this case a polygon with points inside can be circumscribed (Fig. 1). We assume that the compounds located in corners of the polygon adequately represent a family of substances and may be selected as a reference. The software automatically analyses a family of substances and recommends a reference set. The results obtained have shown that this approach for selecting reference compounds allows fairly good predictions (Tables V and VI).

Simulation of chromatographic behaviour and chromatograms

When a column is calibrated under defined conditions, a chromatogram of a sample can be

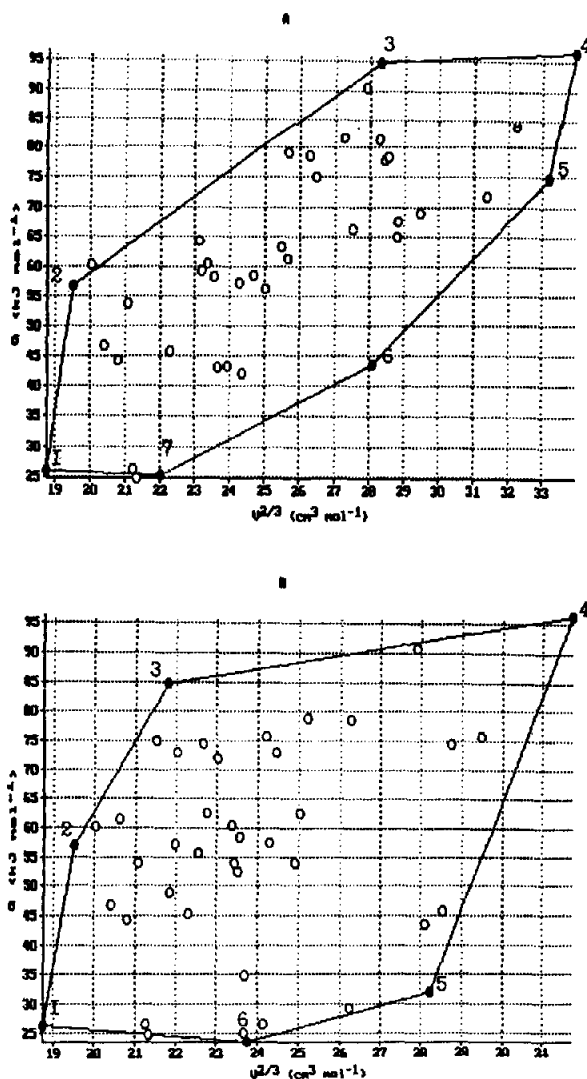
Fig. 1. (A) Selection of reference compounds from 43 compounds (Table VII). 1 = Benzene; 2 = phenol; 3 = *m*-propyl-*p*-hydroxybenzoate; 4 = α-bromo-*p*-phenylacetophenone; 5 = heptanophenone; 6 = biphenyl; 7 = bromobenzene. (B) Selection of reference compounds from 51 compounds (Table VI). 1 = Benzene; 2 = phenol; 3 = *p*-hydroxybenzaldehyde; 4 = diethyl phthalate; 5 = *o*-dichlorobenzene; 6 = butylbenzene.

simulated and a table of predicted k' and ln k' values can be displayed. A fairly large set of calibrated reversed-phase columns (Hypersil ODS, Nucleosil 120-5 C_{18}, Spherisorb ODS-2, Eurospher 80-5 C_{18}, ODS Si-100 Polyol, Separon SGX C_{18}, etc.) and a wide range of organic solvent concentrations (0–100% in steps of 10%)

TABLE V

COMPARISON OF COLUMN PARAMETERS CALCULATED USING DIFFERENT REFERENCE COMPOUNDS

Parameter	43[a]	7[b]	51[c]	6[d]
a	0.313 ± 0.009	0.321 ± 0.007	0.268 ± 0.01	0.279 ± 0.01
b	0.046 ± 0.001	0.046 ± 0.001	0.046 ± 0.001	0.048 ± 0.001
c	-3.775 ± 0.12	-3.783 ± 0.14	-2.569 ± 0.06	-2.698 ± 0.02
R^e	3.17	3.43	4.47	4.65

[a] Reference compounds: 1–43 from Table VII. Column, Spherisorb ODS-2 [5,18]; mobile phase, acetonitrile–water (50:50).
[b] Reference compounds: 5, 12, 15, 17, 30, 36 and 40 from Table VII and Fig. 1A. Column and mobile phase as in footnote a.
[c] Reference compounds: 1–51 from Table VI. Column, Nucleosil C_{18} [19]; mobile phase, methanol–water (40:60).
[d] Reference compounds: 3, 7, 12, 26, 44 and 49 from Table VI and Fig. 1B. Column and mobile phase in footnote c.
[e] $R = \sum_i (\ln k'_{ei} - \ln k'_{ci})^2$.

enables suitable initial conditions to be predicted.

In addition to the possibility of simulating chromatograms and predicting the retention under calibration conditions, interpolation and extrapolation modes can be employed for predicting the retention with different concentrations of an organic solvents in water. When column parameters are determined over a wide range of organic solvent concentrations in a mobile phase, some polynomial can be used to describe the dependences of the column parameters on the concentration of the solvent. Parameters of polynomials have been fitted by the least-squares method from the dependences of column parameters on organic solvent concentration in water ($C_{solv.}$). The software contains a library of cubic polynomial parameters for different ODS columns with methanol–, acetonitrile– and tetrahydrofuran–water mobile phases. These parameters are employed to simulate the dependence of k' (ln k') on $C_{solv.}$, both for a single compound and for mixtures of compounds (up to 100). After drawing and saving several structural formulae in a current library of compounds (Sample), the software calculates partial molar volumes and energies of interaction of the compounds with water and using eqn. 1 and the library of column parameters calculates capacity factors and displays plots of k' (ln k') vs. $C_{solv.}$ for each compound of the Sample (Fig. 2). These dependences enable an acceptable concentration of organic solvent in a mobile phase to be selected. After selection of an organic solvent

concentration, a chromatogram of the mixture on a column with 10 000 plates can be simulated and a table with predicted capacity factor values can be displayed.

Predicting optimum concentration of organic solvent

Some limits of capacity factor values can be considered as optimum depending on the chromatographer's requirements and the sample composition. The software takes two parameters into account: ML, minimum capacity factor value, and MR, maximum capacity factor value. Recommended values are $ML = 1$ and $MR = 10$; however, the limits may be changed by the user. If there is only one compound in a sample and $ML = MR$ the software predicts conditions for $k' = ML = MR$ to be obtained. It seems reasonable to have two modes of optimization, local and global. The local optimization mode considers only one column and one type of organic solvent in a mobile phase. In this case the program predicts the concentration of the organic solvent in water that can provide the maximum correspondence to present ML and MR values. In the global optimization mode the program calculates capacity factors of sample compounds for all columns and mobile phases in the parameters library and recommends conditions (column, type and concentration of organic solvent in water) for acceptable capacity factor values to be obtained. Experimental capacity factor values (range 1.5–5) for a number of compounds listed in refs. 5 and 16 were used to

TABLE VI

COMPARISON OF CALCULATED AND EXPERIMENTAL VALUES OF k'

Column, Nucleosil C_{18} [19]; mobile phase, methanol–water (60:40). V in cm^3 mol^{-1}; $-\Delta G_{e,s,H_2O}$ in kJ mol^{-1}; k'_e = experimental capacity factor values [19]; $k'_{c,1}$ = calculated capacity factor values (column parameters were calculated using reference compounds: 3, 7, 12, 26, 44 and 49); $k'_{c,2}$ = calculated capacity factor values (column parameters were calculated using all 51 compounds as reference).

| No.	Compound	V	$-\Delta G_{e,s,H_2O}$	k'_e [19]	$k'_{c,2}$	$k'_{c,1}$
1	Acetophenone	112.95	60.59	1.90	2.55	2.57
2	Benzaldehyde	96.75	53.80	1.62	1.83	1.88
3	Benzene	81.23	26.16	3.70	3.53	3.49
4	Benzonitrile	92.26	46.73	1.55	2.13	2.16
5	Benzophenone	160.04	75.80	8.04	7.05	6.72
6	Biphenyl	149.04	43.60	25.87	21.92	19.81
7	Butylbenzene	149.99	32.07	45.65	39.11	34.35
8	p-Chlorobenzaldehyde	114.04	52.44	1.48	3.79	3.85
9	Chlorobenzene	98.52	24.80	7.62	7.77	7.43
10	p-Chlorotoluene	115.24	25.03	14.56	14.99	13.85
11	o-Cresol	102.94	57.18	1.67	2.01	2.05
12	o-Dichlorobenzene	115.81	23.44	14.04	16.52	15.19
13	Dimethyl phthalate	147.47	90.66	1.78	2.24	2.26
14	m-Dinitrobenzene	108.47	62.44	2.38	1.96	1.99
15	o-Dinitrobenzene	108.47	62.44	2.03	1.96	1.99
16	p-Dinitrobenzene	108.47	62.44	2.38	1.96	1.99
17	2,4-Dinitrotoluene	125.19	62.67	4.02	3.70	3.66
18	2,6-Dinitrotoluene	125.19	62.67	3.53	3.70	3.66
19	3,4-Dinitrotoluene	125.19	62.67	3.09	3.70	3.66
20	Diphenyl ether	152.44	45.90	23.52	22.19	20.03
21	Ethylbenzene	118.63	26.39	12.23	16.02	14.75
22	m-Fluoronitrobenzene	102.14	48.94	3.53	2.88	2.88
23	o-Fluoronitrobenzene	102.14	48.94	2.49	2.88	2.88
24	p-Fluoronitrobenzene	102.14	48.94	2.48	2.88	2.88
25	p-Fluorophenol	93.51	61.59	1.10	1.11	1.17
26	p-Hydroxybenzaldehyde	101.74	84.59	0.58	0.52	0.57
27	p-Methoxybenzaldehyde	120.87	72.85	1.90	1.91	1.98
28	p-Methylbenzaldehyde	113.47	54.03	2.01	3.45	3.52
29	Naphthalene	115.29	34.88	13.31	9.43	8.91
30	p-Nitroacetophenone	126.57	78.73	1.98	1.83	1.86
31	p-Nitrobenzaldehyde	110.37	71.94	1.33	1.35	1.40
32	p-Nitrilobenzaldehyde	107.78	74.37	0.86	1.10	1.14
33	Nithrobenzene	94.85	44.30	2.72	2.66	2.68
34	m-Nitrophenol	100.34	75.09	1.48	0.76	0.81
35	o-Nitrophenol	100.34	75.09	2.38	0.76	0.81
36	p-Nitrophenol	100.34	75.09	1.34	0.76	0.81
37	p-Phenylphenol	154.03	74.39	5.57	6.10	5.86
38	3-Phenylpropanol	134.76	78.61	2.22	2.40	2.51
39	n-Propylbenzene	134.31	29.23	24.02	25.32	22.76
40	Toluene	97.95	26.39	7.30	7.04	6.76
41	2,4-Dimethylphenol	119.66	57.41	3.06	3.85	3.79
42	2-Phenylethanol	119.08	75.77	1.36	1.58	1.63
43	Benzyl alcohol	103.40	72.93	0.89	0.97	1.03
44	Phenol	86.72	56.95	0.94	1.08	1.07
45	N,N-Dimethylaniline	124.30	53.82	8.12	5.44	5.27
46	N-Methylaniline	108.10	55.26	2.84	2.54	2.56
47	Aniline	89.52	60.30	0.96	1.00	1.05
48	Anisole	105.35	45.21	3.66	3.91	3.86
49	Diethyl phthalate	178.83	96.34	4.74	5.08	4.91
50	Dimethyl phthalate	147.47	90.66	1.78	2.24	2.26
51	Methyl benzoate	114.35	58.41	3.58	2.99	2.98

A

B
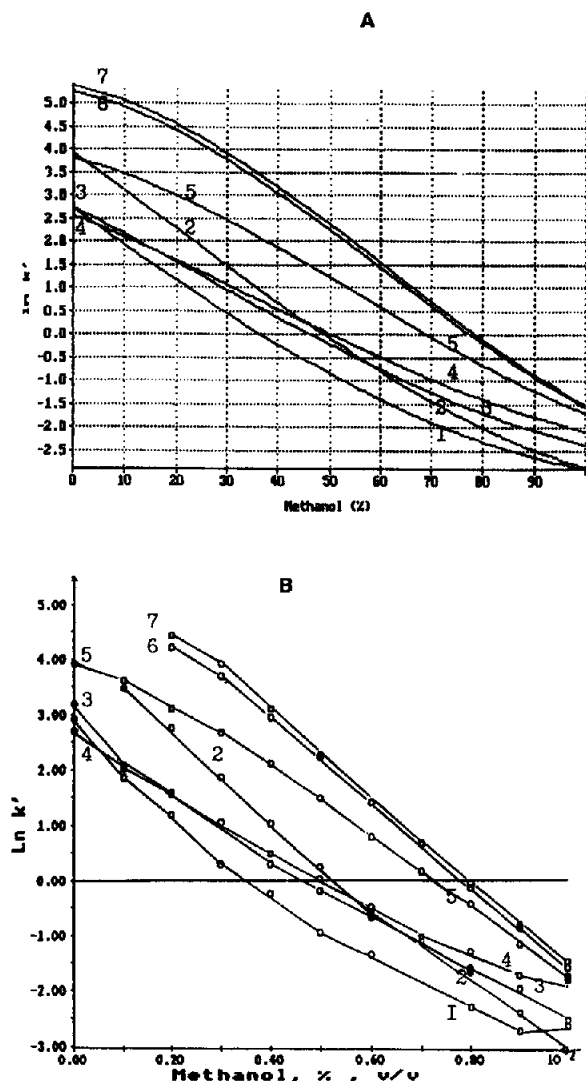
Fig. 2. (A) Simulated and (B) experimental plots of ln k' versus concentration of methanol in water. 1 = p-Hydroxybenzaldehyde; 2 = methylparaben; 3 = phenol; 4 = pyridine; 5 = benzene; 6 = toluene; 7 = chlorobenzene. Sorbent: Nucleosil 120-5 C_{18}; eluent: methanol–water.

check that the program was capable of predicting the concentration of an organic solvent in a mobile phase. The results obtained showed fairly good agreement between experimental and calculated concentrations of an organic solvent in a mobile phase for obtaining required capacity factor values (Table VII).

Isomeric and intramolecular interaction effects

The large set of volume and energies parameters enabled us to predict difference in retention for some isomers. For instance, the program correctly predicts the difference in retention of theobromine and theophylline (Fig. 3). We can request lists of fragments to obtain information on structures analysed (Table VIII).

The software does not take into account stereochemical and intramolecular interaction effects. Sometimes these and other effects can be predicted from a structural formula or from chromatographic experience. For example, capacity factors of *ortho* isomers of benzene derivatives that can form one intramolecular hydrogen bond are nearly twice as large as those for *para* isomers. This value holds fairly well (1.5–2.2) within a wide range of methanol and/or acetonitrile concentrations in a mobile phase [17]. To determine when intramolecular hydrogen bonds can be produced, some considerations should be used. In the simplest approach one can assume that this bond can be formed between H-donors (OH, NH, NH_2, COOH, SH, etc.) and H-acceptors (O- and N-containing groups: NO_2, OH, NH_2, COOH, etc.) if they are in an *ortho*- or α-position to one another, e.g., in *o*-nitrophenol or ethylene glycol one intramolecular H-bond increases the capacity 1.5–2.0-fold. The software allows the user to enter a correction (H_{cor}) for the expected effects. In this case a corrected capacity factor value ($k'_{cor.}$) is calculated as $k'_{cor.} = H_{cor} k'_c$, where k'_c is a capacity factor value calculated without correction. The software takes into account a corrected value both for simulating plots and chromatograms and for predicting the optimum concentration of an organic solvent in a mobile phase. Each compound is saved in the library automatically with $H_{cor} = 1$ (ln $H_{cor} = 0$); however, it may be saved and/or copied with any factors in the range $0.01 < H_{cor} < 2000$.

CONCLUSIONS AND FUTURE DEVELOPMENTS

Software for the prediction of initial conditions in reversed-phase HPLC on the basis of the structural formula of compounds has been developed. The software translates a structural

TABLE VII

EXPERIMENTAL CONCENTRATIONS OF ACETONITRILE [ACN (%) (e)] AND CONCENTRATIONS PREDICTED BY CHROMDREAM [ACN (%) (p)] FOR EXPERIMENTAL CAPACITY FACTORS (k'_e) TO BE OBTAINED

Column, Spherisorb ODS-2 [5,18].

No.	Compound	k'_e	ACN (%) (e)	ACN (%) (p)	Difference (%)	k'_p [a]
1	Acetophenone	2.91	40	45	+5	
2	Aniline	2.21	30	30	0	
3	Anisole	3.43	50	49	−1	
4	Benzaldehyde	3.10	40	38	−2	
5	Benzene	3.42	50	47	−3	
6	Benzonitrile	3.27	40	40	0	
7	Benzyl acetate	2.87	50	50	0	
8	Benzyl alcohol	1.76	30	37	+7	2.42
9	Benzyl bromide	2.84	60	60	0	
10	Benzyl chloride	2.53	60	59	−1	
11	Benzyl cyanide	3.50	40	39	−1	
12	Biphenyl	3.10	70	70	0	
13	1-Bromo-2-nitrobenzene	3.30	50	56	+6	5.20
14	1-Bromo-4-methylphenol	2.38	50	58	+8	4.18
15	α-Bromo-*p*-phenylacetophenone	4.92	60	57	−3	
16	*m*-Bromoaniline	2.14	50	51	+1	
17	*m*-Bromobenzene	3.38	60	59	−1	
18	*p-tert.*-Butylhydroquinone	2.56	60	50	−10	1.35
19	*p-tert.*-Butylphenol	3.50	50	59	+9	7.13
20	Butyrophenone	2.49	60	60	0	
21	Chlorobenzene	3.01	60	59	−1	
22	α-Dibromo-*p*-phenylacetophenone	4.92	60	57	+3	
23	Dimethyl phthalate	1.73	50	54	+4	
24	2,4-Dimethylphenol	1.90	50	58	+8	3.30
25	2,5-Dimethylphenol	2.02	50	57	+7	3.30
26	Ethyl 3-phenylpropionate	3.12	60	63	+3	
27	Ethyl benzoate	2.70	60	56	−4	
28	Ethyl phenylacetate	3.30	50	53	+3	
29	N-Ethylaniline	3.03	50	52	+2	
30	Heptanophenone	2.21	80	79	−1	
31	Hexanophenone	2.90	70	67	−3	
32	Methyl benzoate	2.81	50	50	0	
33	Nitrobenzene	2.43	50	49	−1	
34	4-Nitrophenacyl bromide	2.61	40	40	0	
35	Phenacyl bromide	3.24	50	43	−7	2.00
36	Phenol	2.54	30	28	−2	
37	4-Phenyl-1-butanol	1.95	50	53	+3	
38	5-Phenyl-1-pentanol	2.38	50	56	+6	3.58
39	Propiophenone	2.89	50	52	+2	
40	*m*-Propyl-*p*-hydroxybenzoate	4.33	40	40	0	
41	Thymol	2.46	60	65	+5	
42	Toluene	3.02	60	58	−2	
43	Valerophenone	3.69	60	59	−1	

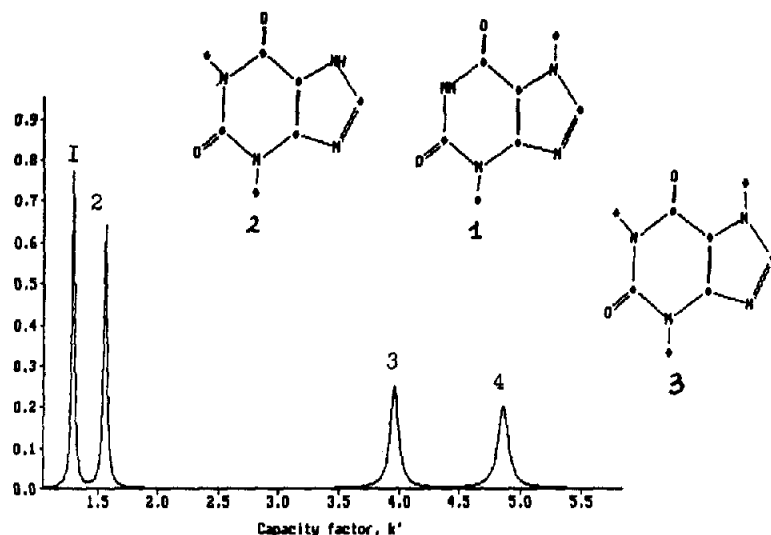[a] When the deviation is >5% or <−5% then predicted values of k' (k'_p) are shown given.

Fig. 3. Simulated chromatogram. 1 = Theobromine; 2 = theophylline; 3 = caffeine; 4 = phenol. Capacity factor range, 1.5–4 for compounds 1, 2 and 3 [20] and 1–10 for compounds 1, 2, 3 and 4 [21]. Sorbent: Nucleosil 120-5 C_{18}; eluent: methanol–water (20:80); column: 13 cm × 4 mm I.D.; flow-rate: 0.5 ml min^{-1}.

TABLE VIII

TRANSLATION OF STRUCTURAL FORMULAE INTO FRAGMENTS

Theobromine		Caffeine		Theophylline	
ar=C–	–2	ar=C–	–2	ar=C–	–2
ar=CH–	–1	ar=CH–	–1	ar=Ch–	–1
CH₃	–2	CH₃	–3	CH₃	–2
ar=N–	–1	ar=N–	–1	ar=N–	–1
arN	–2	arN	–3	arN	–2
arNHp	–1	arCO	–2	arNH	–1
arCO	–2			arCO	–2
V (cm³ mol⁻¹)	117.1		134.4		122.7
ΔG (kJ mol⁻¹)	–123.6		–126.9		–127.8

formula of a compound into fragments and calculates the retention of the compound on different reversed-phase columns with methanol–, acetonitrile– or tetrahydrofuran–water mobile phases.

The software predicts the optimum concentration of methanol, acetonitrile or tetrahydrofuran in water to obtain acceptable values of capacity factors on different reversed-phase columns both for a single compound and for a mixture of compounds (up to 100).

The software simulates chromatograms and plots capacity factors or logarithms of capacity factors as a function of the concentration of an organic solvent in a mobile phase.

The software gives the opportunity to compare different reversed-phase columns, to calibrate columns and to determine parameters of a given column responsible for the retention and selectivity.

We are now working on extensions of the method in several directions: to develop a program for the prediction of the effect of mobile phase pH on retention and to develop a compu-

S.V. Galushko et al. / J. Chromatogr. A 660 (1994) 47–59

59

ter program for the prediction of retention on the basis of structural formulae in gradient reversed-phase HPLC.

ACKNOWLEDGEMENT

REFERENCES

1 T. Hamoir, M. De Smet, H. Phyrins, P. Conty, N.V. Driessche, D.L. Massart, F. Maris, H. Hindriks and P.J. Schoemmakers, *J. Chromatogr.*, 589 (1992) 31.
2 M. De Smet, A. Peeters, L. Buydens and D.L. Massart, *J. Chromatogr.*, 457 (1988) 25.
3 H. Hindriks, F. Maris, J. Vink, A Peeters, M. De Smet and D.L. Massart, *J. Chromatogr.*, 485 (1989) 255.
4 F. Maris, H. Hindriks, J. Vink, A. Peeters, N.V. Driessche and L. Massart, *J. Chromatogr.*, 506 (1990) 211.
5 R. Smith and C.M. Burr, *J. Chromatogr.*, 485 (1989) 325.
6 K. Valko, G. Szabo, J. Rohrich, K. Jemnitz and F. Darvas, *J. Chromatogr.*, 485 (1989) 349.
7 E. Janos, F. Csizmadia, K. Valko and F. Darvas, *19th International Symposium on Chromatography, Aix-en-Provence, 1992*, Abstract 1, p. 58.
8 S.V. Galushko, *J. Chromatogr.*, 552 (1991) 91.
9 S.V. Galushko, *19th International Symposium on Chromatography, Aix-en-Provence, 1992*, Abstract 4, p. 47.
10 S.V. Galushko, *Chromatographia*, 36 (1993) 39.
11 S.V. Galushko, *17th International Symposium on Column Liquid Chromatography, Hamburg, 1993*, Abstract 2-217.
12 S.V. Galushko, *Chromatographia*, in press.
13 S.V. Galushko, A.A. Kamenchuk and G.L. Pit, in preparation.
14 V.P. Belousov and M.Y. Penov, *Termodinamika Vodnyh Rastvorov Neelectrolitov*, Khimiya, Leningrad, 1983, p. 264.
15 B. Bushewski and S.V. Galushko, *Sep. Sci. Technol.*, in press.
16 R.J.M. Vervoort, S.A. Maris and H. Hindriks, *J. Chromatogr.*, 623 (1992) 207.
17 H. Engelhardt and M. Jungheim, *Chromatographia*, 29 (1990) 59.
18 Smith and C.M. Burr, *J. Chromatogr.*, 475 (1989) 57.
19 P.J. Schoemmakers, H.A. Billiet and L. de Galan, *J. Chromatogr.*, 218 (1981) 261.
20 *HPLC Liquid Chromatography 1991; Macherey–Nagel Applications Guide, Application No. 432*, Macherey–Nagel, Düren, 1991.
21 Y. Ohtsu, Y. Shiojima, T. Okumura, J. Koyama, K. Nakamura, O. Nakata, K. Kimata and N. Tanaka, *J. Chromatogr.*, 481 (1989) 147.