Journal of Chromatography A, 660 (1994) 33–46
Elsevier Science B.V., Amsterdam

CHROMSYMP. 2948

Method development by an expert system

Advantages and limitations

Jenő Fekete* and György Morovján

Institute of General and Analytical Chemistry, Technical University of Budapest, Gellért tér 4, H-1111 Budapest (Hungary)

Ferenc Csizmadia and Ferenc Darvas

CompuDrug Chemistry Ltd., P.O. Box 405, H-1395 Budapest 62 (Hungary)

ABSTRACT

Several approaches can be used for the prediction of the optimum eluent composition in RP-HPLC, but only a few are known that use the structure of the solute. The latest release of the computer program EluEx, version 3.0, was developed to help the chromatographer in practical work. The program is based on the prediction of the pK_a and log P (logarithm of 1-octanol–water partition coefficient) values of the solutes. The first eluent suggestion can be done without any preliminary practical work, based on the structural formulae of the solutes. In our experience, two or three experiments are usually sufficient to determine the optimized binary conditions. The surface heterogeneity and the diversity of RP columns, such as the effect of silanol interaction, can be handled by the program only to a limited extent. If the difference in hydrophobicity between two compounds is small, the elution order cannot be predicted properly in all instances. The same is true for some isomers, *e.g.*, diastereomers. In this paper, the results of applying the program to some neutral, acidic and basic solutes are summarized.

INTRODUCTION

In this paper, the results of applying the program EluEx [1], version 3.0, to some neutral, acidic and basic solutes are presented.

According to Mulholland *et al.* [2] and Tsuji *et al.* [3], an expert system embodies the knowledge-based component of an expert's skill in a computer, and the system can offer intelligent advice. Schoenmakers and Dunand [4] defined expert systems in liquid chromatography in a similar way and gave an outline of a complete expert system for method development in HPLC. Hindriks *et al.* [5] have presented an expert system for predicting the initial mobile phase composition., Hamoir *et al.* [6] used the term "first guess" for the selection of initial conditions.

The expert systems created by De Smet *et al.* [7] were based on retention indices introduced to HPLC by Smith [8]. Using retention indices in gas chromatography is a very practical way to predict the relative retentions of different solutes. In liquid chromatography, there are no homologous series such as *n*-alkanes in gas chromatography that are widely accepted for calculating retention indices. The idea of dividing the structure into structural elements and estimating the polarity of analytes has also been utilized in gas chromatography according to Takács *et al.* [9].

In the approach of De Smet *et al.* [7], all initially selected structural elements are related to a percentage of methanol and have a positive or negative value depending on the relative polarity of a structural element. In their earlier publications, a rough estimation was used for predicting the hydrophobicity of analytes, which

* Corresponding author.

0021-9673/94/$07.00
SSDI 0021-9673(93)E0980-9

was obtained from the number of carbon atoms in the investigated molecule. If the carbon number is less than 10, the molecule is considered to be non-hydrophobic, if it is between 10 and 20, the molecule is of medium hydrophobicity, and otherwise it is hydrophobic [10].

De Smet *et al.* [10] used the presence of acidic and basic functional groups in the solutes for the prediction of their acidic and basic character. By applying the expert system for the determination of basic drugs, it was concluded that instead of using structural elements, it seems to be more practical to take the whole chemical structure into account for the prediction of selectivity [11]. To accomplish this, they tried to integrate the DARC system (a computer program for the storage and retrieval of chemicals developed by Télésysteme, Paris, France) with DASH (Drug Analysis System in HPLC), but the programs could only operate independently.

Smith and Burr [12–15] have developed a computer program for the calculation of retention indices and for the prediction of retention, based on molecular structure. They selected benzene as a parent compound and investigated the substituent effect [15]. They found particular problems when the ionized form of the compound was present. In the so-called Prisma model, the retention of compounds in HPLC separations can be predicted by using molecular connectivity indices [16], but no computer-aided method development was presented.

Reviews of expert systems and programs for chromatographic optimization are available elsewhere [17–25] and are outside the scope of this paper. The basics of quantitative structure–chromatographic behaviour can be found in a book by Kaliszan [26].

THEORETICAL BACKGROUND

Problems with the hydrophobicity of the solutes

The similarity of the process of partitioning in 1-octanol–water and RP-HPLC systems makes the prediction of eluent composition possible. A large database of experimental 1-octanol–water partition coefficients [27] and several methods for the prediction of this value can be found in the literature [27–29]. Valkó and co-workers [1,31–35] have described an approach which can be used for the prediction of the appropriate solvent composition from the log P value of the solutes. The basic rule used is the Collander-type rule [30]. The general Collander equation relates the partition coefficients determined in the given organic–water system:

$$\log P_a = m \log P_b + n \tag{1}$$

where P_a and P_b are partition coefficients in solvent systems a and b, respectively, and m and n are constants that are characteristic of the two solvent systems being used. Good agreement can be obtained if the polarity difference between the organic solvents in the two partition systems is small. This suggests that the Collander-type equations can be applied in RP-HPLC if the chromatographic system has similar properties to the 1-octanol–water system [31–35]. The theoretical relationship between reversed-phase retention for non-ionized compounds and the 1-octanol–water partition system has been published by Valkó [34]. Assuming that such a relationship exists, the following equation can be given [31,32]:

$$\log P = m \log K_{chr} + n \tag{2}$$

where K_{chr} is the chromatographic distribution coefficient, which can be related to the capacity factor, k', by

$$k' = \frac{V_s}{V_m} \cdot K_{chr} \tag{3}$$

where V_s/V_m is the phase ratio, which can be treated for practical purposes as a constant value. Substituting eqn. 3 into eqn. 1 and expressing log K_{chr} yields

$$\log P = m \log k' + m \log \left(\frac{V_s}{V_m}\right) + n \tag{4}$$

As $m \log (V_s/V_m) + n$ can be considered to be constant, we arrive at the following equation:

$$\log P = m \log k' + C \tag{5}$$

Eqn. 5 is almost the same as that used by Honai *et al.* [36,37]. It has been shown that retention in RP-HPLC can be approximated well by the

equation [38,39]

$$\log k' = \log k'_w - S\varphi \quad (6)$$

where $\log k'_w$ is the extrapolated value of k' for pure water (or buffer), φ is the volume fraction of the organic component of the solvent and S is a constant, the value of which depends mainly on the solute, but also to some extent on the experimental conditions. Substituting eqn. 6 into eqn. 5, we obtain

$$\log P = m \log k'_w - mS\varphi + m \log \left(\frac{V_m}{V_s}\right) + n \quad (7)$$

Although eqn. 7 is valid, the retention data cannot be predicted using it, because there are two unknown parameters (S and $\log k_w$). Valkó [35] found the following linear relationships:

$$\log P = a\varphi_1 + b \quad (8)$$

where φ_1 is the extrapolated volume fraction of the organic content, when $k' = 1$; φ_1 can be calculated by linear regression from eqn. 6; a and b are constants which can be calculated by linear regression. Rearranging eqn. 8 and applying it to other $\log k'$ values, we obtain

$$\varphi_x = c \log P + d \quad (9)$$

where φ_x is the extrapolated volume fraction when $k' = x$ and c and d are constants. The program EluEx [1] uses two equations with x values of 1 and 5, respectively. To obtain k' values between 1 and 5 we calculate the φ value of the first guess in the following way: φ_1 is calculated for the most hydrophilic (with the lowest $\log P$) and φ_5 for the most hydrophobic (highest $\log P$) compound (φ_1^*, φ_5^*). If $\varphi_5^* < \varphi_1^*$, and $\varphi_5^* < \varphi < \varphi_1^*$ then all k' values should fall in the range 1–5. Hence the following equation is used for the determination of the organic content of the starting eluent:

$$\varphi = \frac{\varphi_1^* + \varphi_5^*}{2} \quad (10)$$

The $\log P$ prediction method used by the program is based on the original work of Rekker and De Kort [40] and further developed by CompuDrug. The structural formula of the compound to be examined is fragmented into groups and interactions among these fragments are taken into account. The structure and the $\log P$ contribution values of the fragments are stored in a database. $\log P$ is calculated as the sum of the contributions of the fragments and the interactions, based on the supposition that substituents cause additive changes in free energy.

$$\log P = \sum_{i=1}^{n} a_i f_i + \sum_{j=1}^{m} b_j F_j \quad (11)$$

where n and m are the number of the type of fragments and interactions, respectively, that occur in the molecule, a_i and b_j are the incidences of fragment i and interaction j, respectively, and f_i and F_j are the $\log P$ contributions of fragment i and interaction j, respectively.

Only the $\log P$ of the non-ionized forms of the compounds are used for the calculation of the organic content, except in the case of ion-pair separations, where the $\log P$ values of the solutes are not directly applicable. The method by which EluEx calculates the eluent composition in ion-pair separations is still under development. This will be the subject of another paper.

Problems of pH determination in mixed aqueous–organic solvents

When secondary equilibrium occurs in a chromatographic system, the given solute may exist in several forms. This affects the distribution between the two phases (*i.e.*, the retention).

If $pH < pK_a - 2$, for an acidic compound, or $pH > pK_a + 2$ for a basic compound, then ionization is almost completely suppressed, which is the aim in most RP-HPLC separations. Because experimental pK_a values are difficult to access, a pK_a prediction module was developed for the current version of the program.

In EluEx the pK_a values are predicted using the Hammett equation for aromatic acids and bases and the Taft equation for aliphatic and alicyclic acids and bases [41]. The assumption of these theories is the same as for $\log P$ prediction, that free energy changes caused by substituents are additive. The form of these equations is

$$pK_a = pK_a^0 - \rho\Sigma\sigma \quad (12)$$

where pK_a^0 is the ionization constant for the parent compound (or protonation reaction cen-

tre), ρ is a constant for the particular equilibrium (characteristic of the centre) and σ is a constant characteristic of a given substituent on a given position for the reaction. EluEx contains a table which stores the pK_a^0 and ρ values of the centres and tables of substituents which store the σ values. The program uses different substituent tables according to the type of centre and the position of the substitution (*ortho*, *meta*, *para*). The centre types are the following:

centres connected to aliphatic parts of the molecule
centres connected to an aromatic system
- anilines
- benzoic acids
- phenols
- pyridines

In the case of condensed aromatic systems, the Dewar–Grisdale method is applied [41]. The distance of the substituent from the centre is also taken into account (*e.g.*, $\sigma_{CH_2Cl} = 0.4\sigma_{Cl}$). The calculation is detailed in a book by Perrin *et al.* [41].

To obtain acceptable k' values and plate numbers for weakly acidic or basic compounds, the dissociation of solutes should be suppressed, or they should be fully ionized. The program can predict the log P of the neutral form of ionizable compounds. Prediction for ionic forms is not yet possible in the current version of the program. As a consequence, the pK_a values must first be predicted, based on the structural formulae. Second, the ion-suppression pH must be derived from the pK_a values. If the program does not accurately estimate the pK_a value, the predicted pH value may be higher than that necessary to suppress the ionization. Because the solute polarity of the ionic species is higher than that of the neutral form, in these instances the retention of the solute investigated may be lower than predicted.

As can be seen, the critical step of the program in these predictions is the accurate estimation of the pK_a values and using these data to determine the pH of the eluent. In a mixed aqueous–organic medium, the pH and pK_a values differ from the values measured in water, and also depend on the ionic strength to some extent.

From the classical Born treatment, a qualitative estimation can be made for the dependence of pK_a on the solvent. In mixed water–organic solvents, the pK_a value can be calculated from

$$\Delta pK_a = \log {}_wK_a - \log {}_sK_a = 122 \cdot \frac{n}{r} \cdot \left(\frac{1}{\epsilon_s} - \frac{1}{\epsilon_w}\right) \quad (13)$$

where w and s refer to water and the given solvent, respectively, r is the ionic radius, ϵ is the dielectric constant and n is a constant which is characteristic of the ionization equilibrium.

For a given solute, $122n/r$ is nearly constant and the variation of the pK_a value with solvent depends on the dielectric constant of water–organic system. The dielectric constants of water, acetonitrile and methanol, frequently used solvents, are 78.5, 37.5, 32.6, respectively at 25°C. The values for acetonitrile and methanol are very close to each other. This means that the same correction can be used for both solvents, because the resulting error is less than that caused by the difference between two reversed phases.

In all instances the pK_a value will increase in water–organic mixtures with increasing organic content. It is known from the literature that a large deviation of pK_a in mixed aqueous solutions from the aqueous pK_a values begins above 80% organic content [42].

The ionic strength has the opposite effect on pK_a according to the Davis equation [43]:

$$\Delta pK_a = 0.512(n-1)I^{1/2} - 0.1(2n-1)I \quad (14)$$

where n = constant (0, 1, 2) and I = ionic strength.

An electrolyte (buffer) makes an acid stronger in the solvatation reaction. The ΔpK_a values depends on the ionization equilibria that occur in aqueous media. Some results were presented by Van de Venne *et al.* [44]. The published ΔpK_a values lie between 0.1 and 0.4. If we compare the solvent effect and the influence of the ionic strength on pK_a, we can conclude that usually there is a decrease in acidity and basicity in mixed water–organic solvents. For example, the pK_a values of benzoic acid is 4.2 in water and 5.0 in methanol–water (40:60, v/v).

To determine the pH of a mixed aqueous–organic solvent is difficult. Even the convention chosen for defining pH is controversial. Schoenmakers *et al.* [45] concluded that there is no generally accepted concept on how to measure the pH in mixed aqueous–organic solvents. If the so-called operational pH (pH*) is used, when the pH is measured by a glass electrode which has been calibrated with an aqueous buffer, then the pH value is usually uncertain [46]. From a thermodynamic point of view, the so-called thermodynamic pH_s should be used, which is based on $p(a_H\gamma_{Cl})$, an acidity function. pH_s will rise with increase in the content of the organic component [46]. The deviation of the pH in the pure aqueous system from pH_s at a 20% methanol content is at least 1 pH unit, at 40% methanol it is at least 2 pH units, and it increases with increasing methanol content [46]. Dissociation phenomena in water–methanol solvents have been widely studied by De Ligny *et al.* [47].

In the early period of liquid chromatography, the importance of pH was realized in the separation of ionogenic solutes [48]. The theoretical treatment of pH in the separation of ionogenic solutes is basically the same as that presented by Horváth *et al.* [48]. Optimization and computer simulation for the prediction of the eluent in the separation of ionogenic compounds has been attracting new interest recently [45,49–53]. Every approach uses the pK_a values of the compounds investigated. Schoenmakers *et al.* [45] measured the pH of the buffer and the retention times of solutes, but they concluded that a physically meaningful pK_a cannot be reliably derived from their model. A software package (Drylab I/mp) was developed for the determination of pK_a from three chromatographic runs, by varying the pH only [50]. For an acceptable determination of pK_a the pH values must be very carefully selected. An approximate value of pK_a is needed before the experiments.

When separating ionogenic compounds, the selectivity is the highest around the pK_a value, but Schoenmakers *et al.* [45] showed that the plate number and the asymmetry factor varied for nitrophenols. When the ionization was suppressed, higher plate numbers and better peak shapes were observed. In general, if two equilibria take place simultaneously, the peak will broaden and will be asymmetric. If the dissociation of the analyte is faster than its distribution process, then only one peak will appear on the chromatogram.

Rules for calculating the pH of the buffer

The sample may contain compounds with different acidic and basic characteristics. EluEx places every sample into one of the classes listed below. Because acidity and basicity are lower in a mixed aqueous–organic medium than in water, the following correction on the pK_a values was used:

$$pK_a^* = pK_a + a\varphi \quad (15)$$

where $a = 1$ for acids and -2 for bases. Eqn. 16 is supposed to consider both pK_a and pH changes in a mixed aqueous–organic eluent compared with pure water (or buffer). The replacement of this correction by a more theoretical approach, which is based on a pH and pK_a prediction in a mixed aqueous environment, is under development.

The following notations are used:

$pK_a^{a^*}$	the lowest corrected acidic pK_a of the strongest acid in the sample
$pK_a^{b^*}$	the highest corrected basic pK_a of the strongest base in the sample
NA	none of the solutes are acidic or $pK_a^{a^*} > 11$ (that is, the acidities of all solutes are negligible)
NB	none of the solutes are basic or $pK_a^{b^*} < 1$ (the basicities of all solutes are negligible)
min pH, max pH	set by the user, determined by the type of the column and the solutes
C	condition
A	action
E	explanation

(a)	C	NA and NB (neutral compounds or very weak acids or bases)
	A	buffer is not suggested
	E	the acidities and basicities of the solutes are so weak, if they exist, that they will not be ionized in a buffer-free eluent

(b) C min $pH + 2 < pK_a^{a*} \leq 11$ and NB (weakly acidic compounds)
A $pH = pK_a^{a*} - 2$
E ionization is suppressed if $pH \leq pK_a^{a*} - 2$, because $[A]/[B] \geq 100$ (see eqn. 21).

(c) C NA and $1 \leq pK_a^{b*} < \max pH - 2$ (weakly basic compounds)
A $pH = pK_a^{b*} + 2$, if $pK_a^{b*} > 4$ then a masking agent (*e.g.*, triethylamine) against the silanol effect is suggested
E ionization is suppressed if $pH \geq pK_a^{b*} + 2$, because $[B]/[A] \geq 100$.

(d) C $pK_a^{a*} \leq \min pH + 2$ and NB (acidic compounds, ionization cannot be suppressed in the given pH range)
A pH is set to be the furthest from all pK_a values of the acids in the pH range. Basic ion-pair reagent is suggested
E acids with $pK_a < pH$ will form ion pairs with the reagent, the dissociation of the other compounds will be suppressed

(e) C NA, $pK_a^{b*} \geq \max pH - 2$ (basic compounds, ionization cannot be suppressed in the given pH range)
A pH is set to be the furthest from all pK_a values of the bases in the pH range. Acidic ion-pair reagent is suggested.
E bases with $pK_a > pH$ will form ion pairs with the reagent, the dissociation of the other compounds will be suppressed

(f) C $pK_a^{a*} \leq 11$, $pK_a^{b*} \geq 1$, $pK_a^{a*} > pK_a^{b*} + 4$ (weak acids and weak bases)
A $pH = (pK_a^{a*} + pK_a^{b*})/2$, if $pK_a^{b*} > 4$ then a masking agent against the silanol effect is suggested
E if $pK_a^{b*} + 2 \leq pH \leq pK_a^{a*} - 2$ then the ionization of all compounds is suppressed

(g) C $pK_a^{a*} \leq 11$, $pK_a^{b*} \geq 1$, $pK_a^{a*} \leq pK_a^{b*} + 4$ and $pK_a^{a*} - \min pH < \max pH - pK_a^{b*}$ (acids and weak bases)
A pH is set to be the furthest from all pK_a values of the acids in the range $[\max(pK_a^{b*} + 2, pK_a^{a*} + 2), \max pH]$ if possible, otherwise max pH. Basic ion-pair reagent is suggested
E Acids and $pK_a < pH$ will form ion pairs with the reagent, the dissociation of the other compounds is assumed to be suppressed.

(h) C $pK_a^{a*} \leq 11$, $pK_a^{b*} \geq 1$, $pK_a^{a*} \leq pK_a^{b*} + 4$ and $pK_a^{a*} - \min pH \geq \max pH - pK_a^{b*}$ (weak acids and bases)
A pH is set to be the furthest from all pK_a values of the bases in the range $[\min pH, \min(pK_a^{a*} - 2, pK_a^{b*} - 2)]$ if possible, otherwise min pH. Basic ion-pair reagent is suggested
E bases with $pK_a > pH$ will form ion pairs with the reagent, the dissociation of the other compounds is assumed to be suppressed

There may be cases when there are both strong acids and strong bases in the sample. This version of the program cannot handle that situation, but it is under development.

Isocratic optimization

The optimization method of the program is based on the window diagram technique. The organic content is varied to obtain a model for the system. After two experiments, eqn. 7 is used for the k' prediction. The parameters log k_w and S are calculated from the data for the two chromatograms by linear regression. After three or more experiments a parabolic model is used for modelling the system's behaviour between the lowest and highest experimental φ values. Outside this range the curve's tangents for these two points are applied.

The goal of optimization is to obtain the fastest run with an $R_{s,min}$ value higher than or equal to the user defined value, with peaks in the k' range set by the user.

The optimization is completed when the difference between the last and the previous organic

content suggestion is less than 1%. Of course, users can stop at any time when they find the chromatogram to be acceptable.

Gradient optimization

In cases when isocratic elution could be completed but the resolution is poor or the elution time is long, it may be reasonable to run gradient elution. The simulation of LSS (linear solvent strength) gradient runs is possible using the data for two isocratic runs. The equations used for this prediction can be found in ref. 54.

The equation for the calculation of retention time in gradient elution is

$$t_g = \frac{t_0}{b} \cdot \log (2.3k_a b + 1) + t_0 \quad (16)$$

where k_a is the isocratic k' value with the starting eluent and b is the gradient steepness, which can be derived from the following equation:

$$b = \frac{t_0}{t_G} \cdot (\log k_a - \log k_b) \quad (17)$$

where k_b is the k' value using isocratic elution with the final eluent composition of the gradient; k_a and k_b can be calculated from eqn. 6.

The optimization criteria are the same as in isocratic elution. For LSS gradient elution the definition of R_s is the same as in the isocratic case, but we have to predict the band width of compound i, W_i, with the following equation:

$$W_i = \frac{G\left(1 + \frac{k_a}{2.3k_a b + 1}\right)t_0}{N^{1/2}} \quad (18)$$

where G is the band compression factor, which can be calculated from

$$G^2 = \frac{1 + p + p^2}{(1 + p)^2} \quad (19)$$

where p is given by

$$p = \frac{2.3k_a b}{k_a + 1} \quad (20)$$

Structure of the program

EluEx is written in C and C++ languages and runs on IBM AT-compatible computers. The program is built from several modules:

(1) The main module controls the activities of the other modules.

(2) In the structure maintenance module the structural formulae of the solutes can be entered into the compound database. The structures can be drawn easily using a mouse.

(3) The log P and pK_a prediction modules calculate these physico-chemical parameters from the structural formulae.

(4) The initial step module calculates the initial eluent composition, that is, the organic percentage, and if needed the pH of the buffer and the amount and type of the ion-pair agent or masking agent.

(5) The optimization module governs the optimization phase, which is made up of experiments and new suggestions which use the data obtained from the previous experiments.

(6) The simulation module can simulate the chromatogram graphically for varying organic percentages.

Working with the program

Fig. 1 shows the flowchart of the program. The following steps have to be made during a session:

(1) Entering the structures with a mouse, if they are not already in the compound database.

(2) Composing the list of solutes by selecting the compounds from the database.

(3) Defining:
- the plate number;
- the preffered k' range;
- the chosen organic modifier (six are available in the program);
- the chosen acidic or basic ion-pair reagent (six and five are available, respectively);
- the pH range; and
- $R_{s,min}$.

(4) As one enters the initial step module, it calculates the first-step eluent composition.

First the log P and pK_a values of the solutes are predicted using fragmentation methods.

The type of the sample is determined using the smallest acidic and the greatest basic pK_a values of the set of compounds.

If max log P − min log $P > 5$, gradient elution will be suggested.

The organic percentage will be calculated from the smallest and highest log P values.

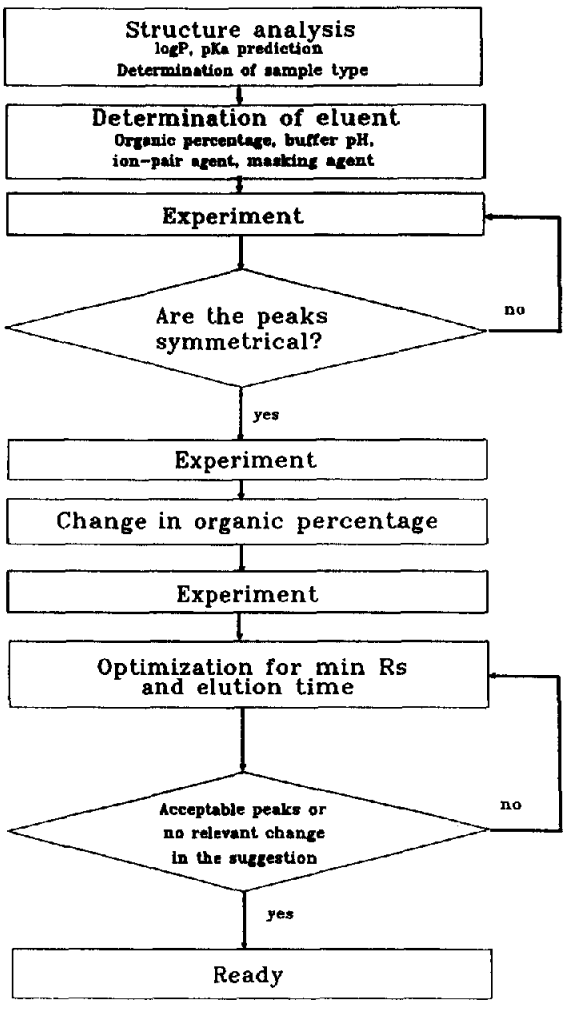

Fig. 1. Flow chart of the program.

(5) The user performs an experiment using the suggested eluent and enters the data into the program. If matrix components exist, their retention data should be entered, so that they too will be considered.

(6) As long as the peaks of interest are not symmetrical, modifications to the eluent composition will be suggested, such as changing pH to max pH in the case of bases or min pH in the case of acids, or increasing the amount of the ion-pair or the masking agent.

(7) If the peaks are symmetrical, the organic percent is changed to collect data for the modelling of the system. Again, following the experiment the data must be entered.

(8) At this point, as we have two φ–log k' data points for all solutes, the behaviour of the system can be modelled using eqn. 5. This equation is used for both the calculation of the optimized organic percentage and the chromatogram simulation.

(9) If the criteria for k' range and R_s can only be satisfied with gradient elution, an optimized gradient profile is suggested and the optimization process is halted.

(10) After each iosocratic run the model is refined (parabolic instead of linear) to obtain more exact k' predictions for the optimization and the simulation.

EXPERIMENTAL

Chromatographic apparatus

Two types of HPLC apparatus were used for the experiments: (1) Millipore (Milford, MA, USA) Model 510 isocratic pump and a Millipore DAD 911 diode-array detector and (2) an HP1090 system equipped with a Model 1040 UV–Vis detector and a Chemstat data station with HP7958B hardware and software (Hewlett-Packard, Rockville, MD, USA).

Chemicals and reagents

The solvents were of chromatographic grade. Methanol and acetonitrile were obtained from Merck (Darmstadt, Germany). Deionized water as purified by means of a Milli-Q system (Millipore). Other chemicals were of analytical-reagent grade from Merck. Some triazines were purchased from Supelco (Gland, Switzerland) and others were donated by the Hungarian National Plant Protection Institute. Fumagillin, mevinoline and mevinolinic acid were produced and purified at the Institute of General and Analytical Chemistry, Technical University of Budapest.

The column for chlorophenol, fumagillin and mevinolines was packed with LiChrosorb RP-18,10 μm (Merck) (250 × 4.6 mm I.D.) and for triazines with Supelcosil LC-18-DB (Supelco) (150 × 4.6 mm I.D.) and Spherisorb ODS-2,5 μm (Hewlett-Packard) (250 × 4.6 mm I.D.).

The pH was adjusted using a Horiba F-8 pH meter.

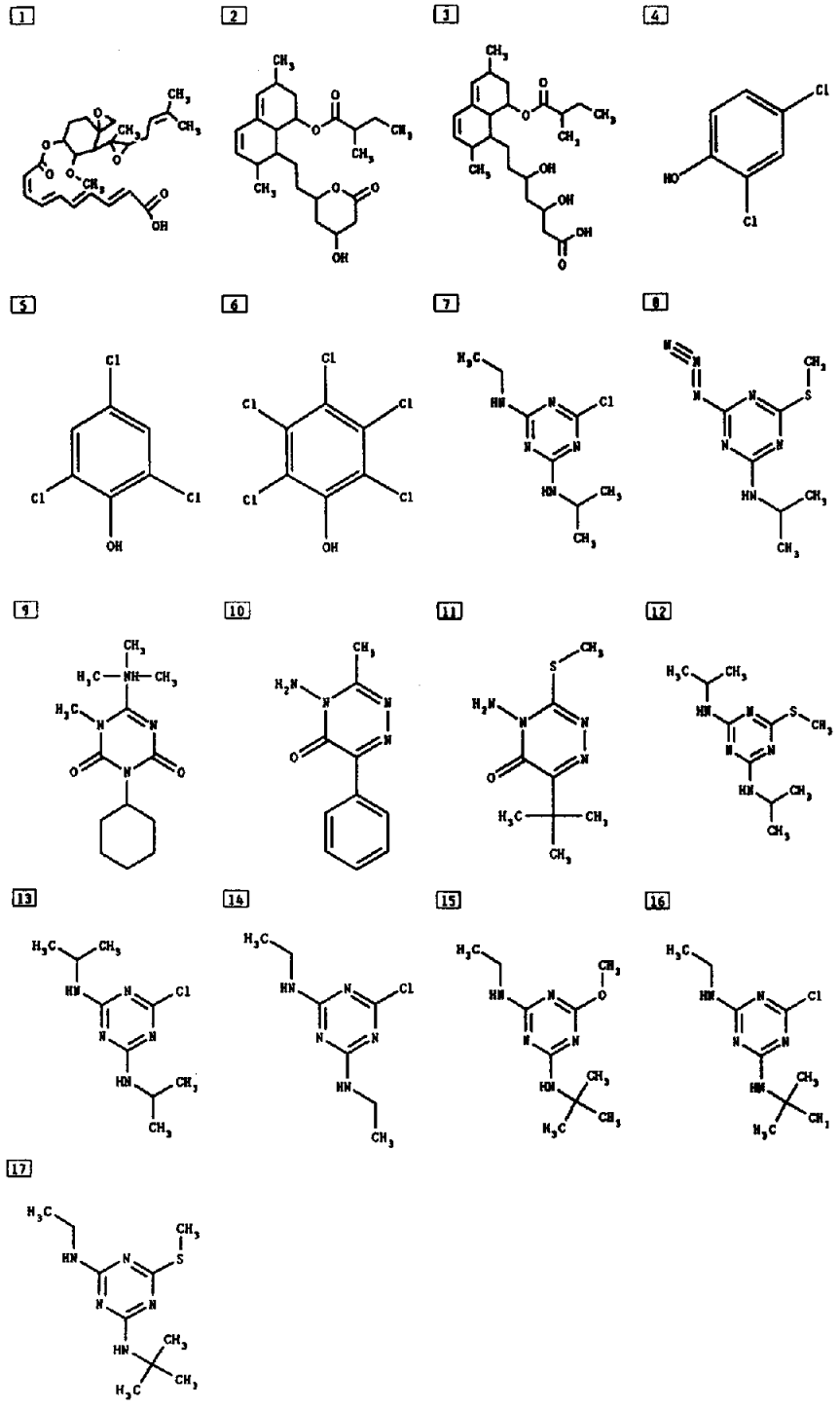

Fig. 2. Structures of solutes investigated. **1** = fumagillin; **2** = mevinoline; **3** = mevinolinic acid; **4** = 2,4-dichlorophenol; **5** = 2,4,6-trichlorophenol; **6** = pentachlorophenol; **7** = aziprotryne; **8** = metamitron; **9** = hexazinon; **10** = metribuzin; **11** = simazine; **12** = terbumeton; **13** = atrazine; **14** = terbutryn; **15** = prometryn; **16** = terbutilazine; **17** = propazine.

RESULTS

Determination of fumagillin in biological matrix

The separation of fumagillin (compound **1** in Fig. 2) from a fermentation liquid is considered. The predicted pK_a and log P values of fumagillin are 3.2 and 4.76, respectively. The suggested eluent composition in the first guess is 90.0% (v/v) acetonitrile–50 mM KH_2PO_4 buffer (pH 2.1). The measured value for k' is 0.66 (Fig. 3A). The matrix components were eluted near fumagillin. In the second guess, 75% (v/v) acetonitrile was suggested. The k' value of fumagillin became 1.29 and the separation was still not complete (Fig. 3B). In the third step, the program suggested 60% (v/v) acetonitrile. Using this composition, the separation became perfect and the k' of fumagillin was 3.45 (Fig. 3C).

Determination of mevinolines in biological matrix

The structures of mevinoline and mevinolinic acid are shown in Fig. 2 (compounds **2** and **3**). The program suggested 97.5% (v/v) methanol–

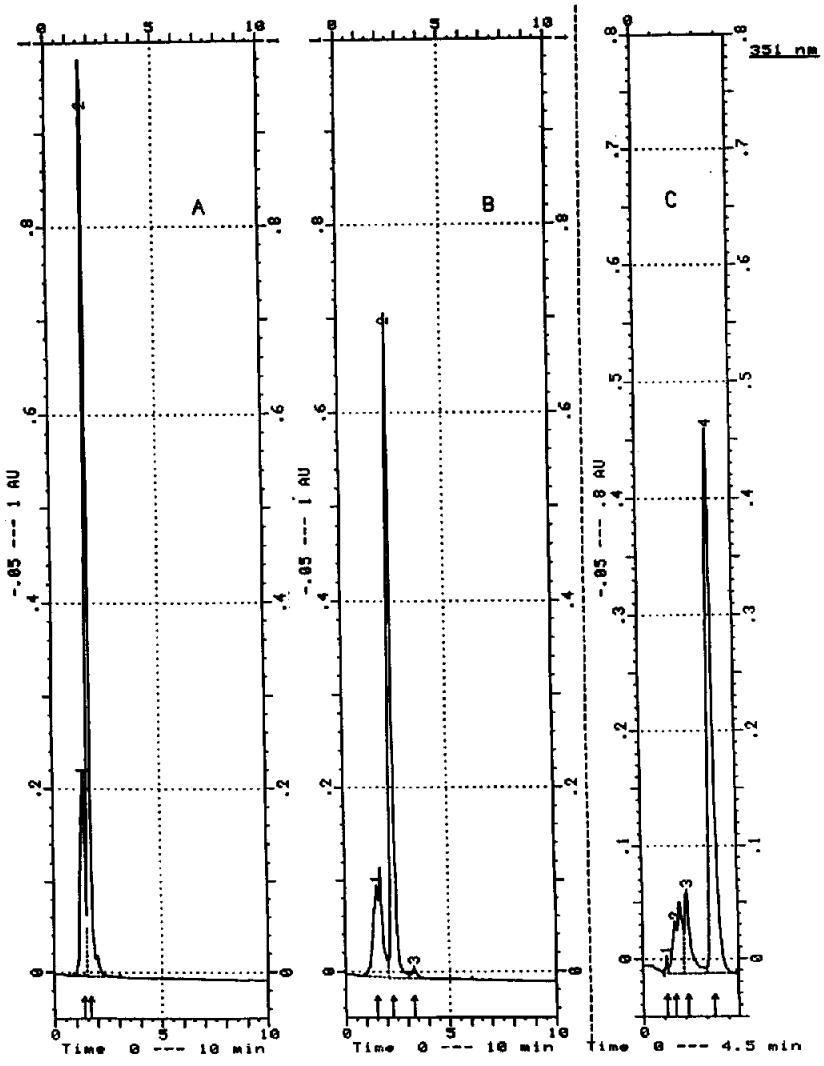

Fig. 3. Determination of fumagillin in fermentation liquid. Column, LiChrosorb RP-18, 10 μm (250 × 4.6 mm I.D.); buffer, 50 mM KH_2PO_4 (pH 2.1); eluent, (A) 90, (B) 75 and (C) 60% (v/v) acetonitrile–buffer; flow-rate, 1 ml/min.

m*M* KH_2PO_4 buffer (pH 3.3). There was no acceptable separation between the matrix components and the solutes investigated. In the second step, the program suggested 80% (v/v) methanol. Completing the experiment with this eluent, we obtained an acceptable separation (Fig. 4).

Determination of chlorophenols

The test compounds with pK_a values between 4 and 9 were 2,4-dichlorophenol,3,4,6-trichlorophenol and pentachlorophenol (compounds **4**, **5** and **6** in Fig. 2). The calculated pK_a values were 7.9, 6.3 and 4.7, respectively, and the log *P* values were 2.98, 3.72 and 5.20, respectively. The suggested eluent composition for the first trial was 79% (v/v) acetonitrile–50 m*M* KH_2PO_4 buffer (pH 3.5).

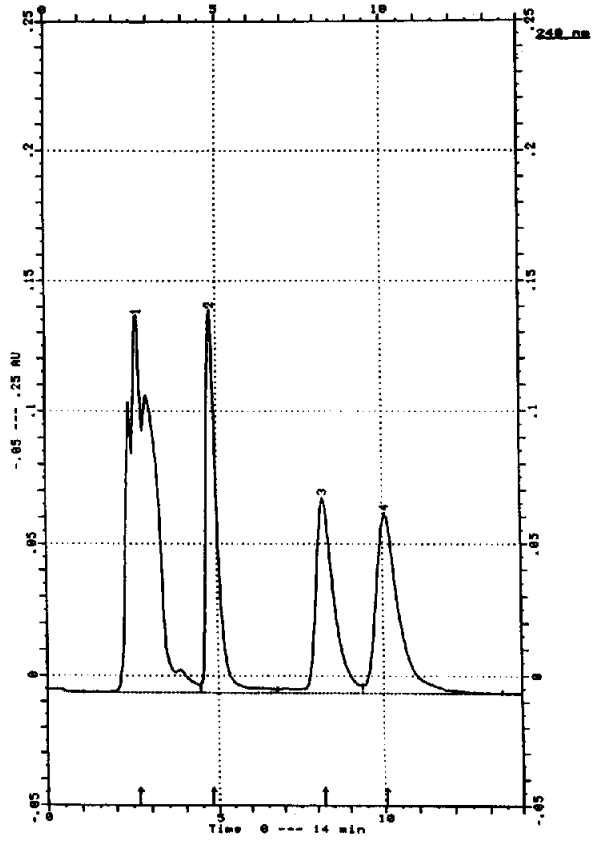

Fig. 4. Determination of mevinolines in fermentation liquid. Column, LiChrosorb RP-18, 10 μm (250 × 4.6 mm I.D.); buffer, 50 m*M* KH_2PO_4 (pH 3.3); eluent, 80% (v/v) methanol–buffer; flow-rate, 1 ml/min. Peaks: 1 and 2 = matrix components; 3 = mevinolinic acid; 4 = mevinoline.

The capacity factors of the first compounds using these conditions [with 80% (v/v) acetonitrile] were 0.24, 0.40 and 0.80, respectively (Fig. 5A). As the peaks were symmetrical in the second guess, the suggested mobile phase composition differed only in the organic content. The new eluent suggestion was 50% (v/v) acetonitrile. On carrying out the measurement the capacity factors of the test solutes were 1.49, 2.59 and 6.21, respectively (Fig. 5B). The chromatogram was then acceptable, so the method development was completed.

Determination of some triazine herbicides

Triazine herbicides were used to test the program for weakly basic solutes. The pK_a values of triazine compounds are relatively small (pK_a < 4.5) [55]. Even at this small pK_a value, the silanophilic interaction can cause a large decrease in the efficiency of the separation and results in bad peak shapes. The structures of triazine compounds are given in Fig. 2 (compounds **7–17**).

The program suggested 39% (v/v) acetonitrile–50 m*M* KH_2PO_4 buffer (pH 5.5). First we used 40% (v/v) acetonitrile–buffer. In a conventional column (Spherisorb ODS-2) the separation was acceptable (Fig. 6A). The peak symmetry was fairly good except for azyprotryn. To diminish the effect of silanophilic interaction, a deactivated RP column was used (Supelcosil 8DB). The chromatographic separation is shown in Fig. 6B. All components are separated. In Table I the calculated log *P* and pK_a values are given. As can be seen from these values, the experimental retention order does not correspond to the predicted hydrophobicity order in some instances.

The eluent composition was predicted with methanol to check the selectivity. The program suggested 59% (v/v) methanol, so the solutes were separated at 60% (v/v) methanol. The chromatogram is shown in Fig. 6C. As can be seen, using methanol we did not obtain an acceptable separation.

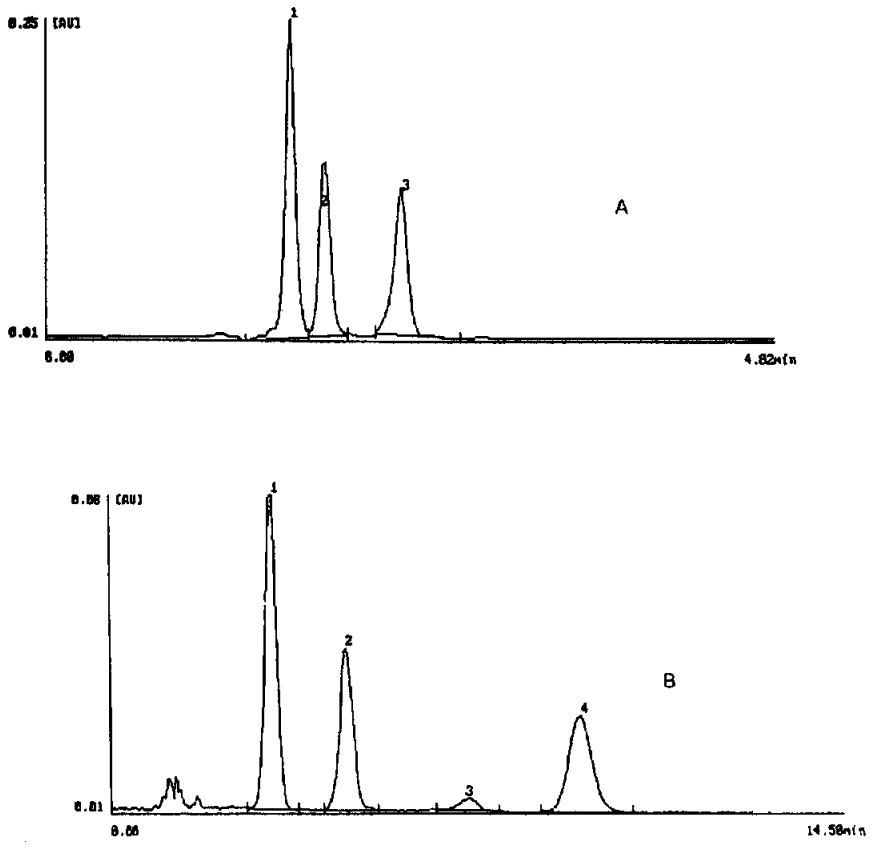

Fig. 5. Separation of chlorophenols. Column, LiChrosorb RP-18, 10 μm (250 × 4.6 mm I.D.); buffer, 50 m*M* KH_2PO_4 (pH 3.5); eluent, (A) 80 and (B) 50% (v/v) acetonitrile–buffer; flow-rate, 1 ml/min. Peaks: (A) 1 = 2,4-dichlorophenol; 2 = 2,4,6-trichlorophenol; 3 = pentachlorophenol; (B) 1 = 2,4-dichlorophenol; 2 = 2,4,6-trichlorophenol; 3 = impurity; 4 = pentachlorophenol.

TABLE I

HYDROPHOBICITY ORDER OF TRIAZINE HERBICIDES

Compound	Log *P*	pK_a
Aziprotryne	−0.41	1.4
Metamitron	0.88	<1
Hexazinon	1.21	<1
Metribuzin	1.65	<1
Simazine	2.24	1.4
Terbumeton	2.61	3.7
Atrazine	2.76	1.2
Terbutryn	3.14	4.1
Prometryn	3.14	4.3
Terbutilazine	3.28	1.2
Propazine	3.28	1.4

CONCLUSIONS

The program EluEx was tested for the prediction of eluent composition with neutral, weakly acidic and weakly basic compounds. The program predicts the distribution constant of the non-ionized form of the solutes from hydrophobic fragmental values. Assuming a linear relationship between log *P* and log *k'*, the organic content can be predicted from the structural formulae. For acidic and basic solutes the program predicts pK_a values based on Hammett and Taft equations. For weak acids and bases the pK_a values are used for the prediction of buffer pH. It was shown that in most instances during the initial step the program can predict an eluent

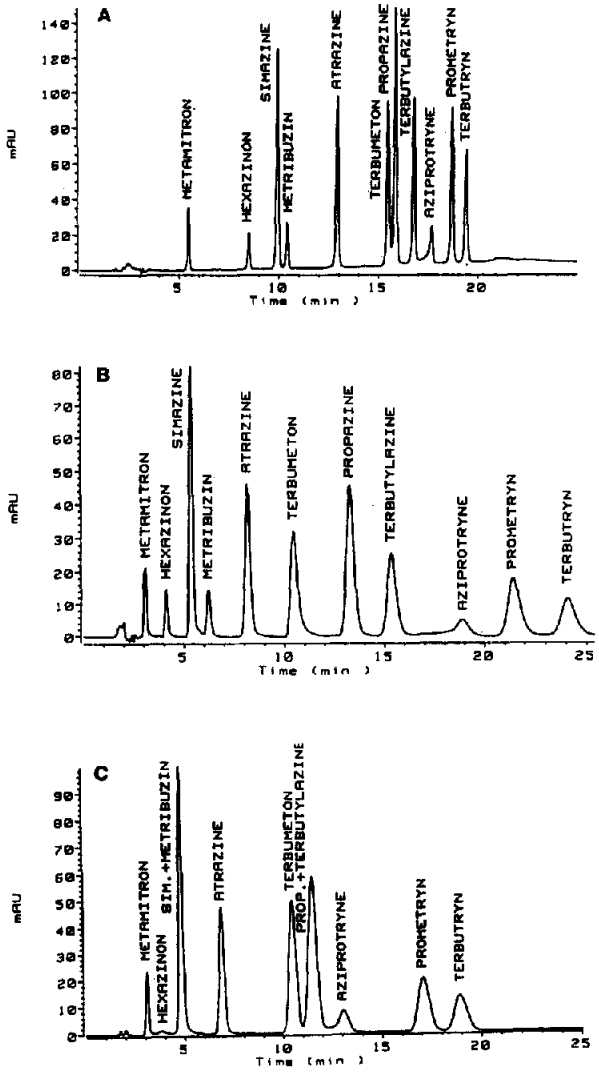

Fig. 6. Separation of triazine herbicides. Column, (A) Spherisorb ODS, 5 μm (250 × 4.6 mm I.D.) (B and C) Supelcosil LC-18-DB (150 × 4.6 mm I.D.); buffer, 50 m*M* KH_2PO_4 (pH 5.5) (adjusted with 1 *M* NaOH); eluent, (A and B) 40% (v/v) acetonitrile–buffer and (C) 60% (v/v) methanol–buffer; flow-rate, 0.8 ml/min.

composition that is a good starting point for the method development. By entering the experimental data as needed and following the suggestions of the program, the method development can be completed usually by the second or third experiment. In some instances the elution order differs from the predicted hydrophobicity order. The correction of the predicted pK_a value for mixed aqueous solvents has to be developed further.

SYMBOLS

pK_a negative logarithm of the acid dissociation constant:

$$pK_a = pH + \log \frac{[A]}{[B]} \quad (21)$$

where A is the acidic and B is the basic form of a compound (in the case of bases the pK_a value of the protonated compound as an acid is used as the pK_a).

Log *P* logarithm of the partition coefficient of a compound between 1-octanol and water.

t_0 column dead time (min).

t_R solute retention time (min).

k' capacity factor for a given band, equal to $(t_R - t_0)/t_0$.

R_s resolution, equal to $2(t_2 - t_1)/(W_1 + W_2)$, where t_1 and t are retention times $(t_2 > t_1)$ and W_1 and W_2 are their baseline band widths.

$R_{s,min}$ resolution of the most poorly resolved band pair.

φ volume fraction of organic solvent in a binary mobile phase.

ACKNOWLEDGEMENTS

This research was supported by the OTKA (contract number 117), SZTO and OMFB foundations. We acknowledge the assistance of Robert J. Longworth, Szilvia Kokai, Ildikó Szepesi and Csilla Sárközi.

REFERENCES

1 P. Csokán, K. Valkó, F. Darvas and F. Csizmadia, *LC·GC*, 6 (1993) 361.
2 M. Mulholland, N. Dunand, A. Cleland, J.A. van Leeuwen and B.G.M. Vandeginste, *J. Chromatogr.*, 485 (1989) 293.
3 K. Tsuji and K.M. Jenkins, *J. Chromatogr.*, 485 (1989) 297.
4 P.J. Schoenmakers and N. Dunand, *J. Chromatogr.*, 485 (1989) 219.

5 R. Hindriks, F. Maris, J. Vink, A. Peters and M. De Smet, D.L. Massart, *J. Chromatogr.*, 485 (1989) 255.
6 T. Hamoir, M. de Smet, H. Pyrins, P. Conti, N. Vandendriesche, D.L. Massart, F. Morris, H. Hindriks and P.J. Schoenmakers, *J. Chromatogr.*, 589 (1992) 31.
7 M. De Smet, G. Musch, A. Peeters, L. Buydens and D.L. Massart, *J. Chromatogr.*, 485 (1989) 237.
8 R.M. Smith, *Adv. Chromatogr.*, 26 (1987) 277.
9 J. Takács, C. Szita and G. Tarján, *J. Chromatogr.*, 56 (1971) 1.
10 M. De Smet, A. Peeters, L. Buydens and D.L. Massart, *J. Chromatogr.*, 457 (1988) 25.
11 F. Maris, R. Hindriks, J. Vink, A. Peeters, N.V. Priessche and L. Massart, *J. Chromatogr.*, 506 (1990) 211.
12 R.M. Smith and C.M. Burr, *J. Chromatogr.*, 465 (1989) 75.
13 R.M. Smith and C.M. Burr, *J. Chromatogr.*, 481 (1989) 71.
14 R.M. Smith and C.M. Burr, *J. Chromatogr.*, 485 (1989) 325.
15 R.M. Smith and C.M. Burr, *J. Chromatogr.*, 550 (1991) 335.
16 Sz. Nyiredi, *Application of the Prisma Model for the Selection of Eluent Systems in Overpressure Layer Chromatography (OPLC)*, Labor MIM, Budapest, 1987.
17 P.J. Schoenmakers, *Optimization of Chromatographic Sensitivity –Guide to Method Development*, Elsevier, Amsterdam, 1986.
18 J.C. Berridge, *Techniques for the Automated Optimization of HPLC Separations*, Wiley-Interscience, New York, 1989.
19 L.R. Snyder, J.L. Glajch and J.J. Kirkland, *Practical HPLC Development*, Wiley-Interscience, New York, 1989.
20 R.L. Glajch and L.R. Snyder (Editors), *Computer-Assisted Method Development for HPLC*, Elsevier, Amsterdam, 1990.
21 A. Peeters, L. Buydens, D.L. Massart and P.J. Schoenmakers, *Chromatographia*, 26 (1988) 101.
22 D. Goulder, T. Blaffert, A. Blokland, L. Buydens, A. Chhabra, A. Cleland, N. Dunand, H. Hindriks, G. Kateman, H. Van Leeuwen, D.L. Massart, M. Mulholland, G. Musch, P. Naish, A. Peeters, G. Postma, P. Schoenmakers, M. De Smet, B. Vandeginste and J. Vink, *Chromatographia*, 26 (1988) 237.
23 P.J. Schoenmakers, A. Peeters and R.J. Lynch, *J. Chromatogr.*, 506 (1990) 169.
24 J.W. Dolan, D.C. Lommen and L.R. Snyder, *J. Chromatogr.*, 535 (1990) 55.
25 L.R. Snyder, J.W. Dolan and D.C. Lommen, *J. Chromatogr.*, 535 (1990) 75.
26 R. Kaliszan, *Quantitative Structure–Chromatographic Retention Relationships*, Wiley, New York, 1987.
27 C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York, 1979.
28 Han van de Waterbeemd and F. Darvas (Editors), *Hydrophobicity of Organic Compounds*, CompuDrug, Vienna, 1986.
29 C. Hansch and A. Leo, *Log P and Parameter Database: a Tool for the Quantitative Prediction of Bioactivity*, Pomona College, Medicinal Chemistry Project, 1983, issue 22,23.
30 R. Collander, *Acta Chem. Scand.*, 5 (1951) 774.
31 G. Szepesi and K. Valkó, *J. Chromatogr.*, 550 (1991) 7.
32 K. Valkó and P. Segel, *J. Chromatogr.*, 631 (1993) 49.
33 K. Valkó, *J. Liq. Chromatogr.*, 10 (1987) 1663.
34 K. Valkó, *J. Liq. Chromatogr.*, 7 (1984) 1405.
35 K. Valkó, *Microcomputer Analysis of HPLC, the State of the Art*, Akadémiai Kiadó, Budapest, 1985, pp. 739–750.
36 T. Honai, *Chromatographia*, 12 (1979) 77.
37 T. Honai, K.C. Tran and J. Hubert, *J. High. Resolut. Chromatogr. Chromatogr. Commun.*, 4 (1981) 454.
38 W.R. Melander and Cs. Horváth, in Cs. Horváth (Editor), *High-Performance Liquid Chromatography –Advances and Perspectives*, Vol. 2, Academic Press, New York, 1980, p. 114.
39 P. Jandera and J. Churáček, *Gradient Elution in Column Liquid Chromatography*, Elsevier, Amsterdam, 1985, p. 19.
40 R.F. Rekker and H.M. de Kort, *Eur. J. Med. Chem.*, 14 (1979) 479.
41 D.D. Perrin, B. Dempsey and E.P. Serjeant, *pK_a Prediction for Organic Acids and Bases*, Chapman and Hall, London, 1981.
42 A.L. Bacarella, E. Gunwald, H.P. Marshall and E.L. Purlee, *J. Org. Chem.*, 20 (1955) 747.
43 C.W. Davies, *J. Chem. Soc.*, (1938) 2039.
44 L.L.M. van de Venne, I.L. Hendriks and R.S. Deelder, *J. Chromatogr.*, 167 (1978) 1.
45 P.J. Schoenmakers, S. van Molle, C.M.G. Hayes and L.G.M. Uunk, *Anal. Chim. Acta*, 250 (1991) 1.
46 R.G. Bates, *Determination of pH. Theory and Practice*, Wiley, New York, 1973, Ch. 8.
47 C.L. de Ligny, H. Loriaux and A. Ruitar, *Recl. Trav. Chim. Pays-Bas*, 80 (1961) 725.
48 Cs. Horváth, W. Melander and I. Molnár, *Anal. Chem.*, 49 (1977) 142.
49 R.M.L. Marques and P.J. Schoenmakers, *J. Chromatogr.*, 592 (1992) 157.
50 J.A. Lewis, D.C. Lommen, W.D. Raddatz, J.W. Dolan and L.R. Snyder, *J. Chromatogr.*, 592 (1992) 183.
51 L.J.W. Dolan and L.R. Snyder, *J. Chromatogr.*, 592 (1992) 183.
52 P.J. Schoenmakers, N. Mackie and R.M.L. Marques, *Chromatographia*, 35 (1993) 18.
53 R.M.L. Marques, P.J. Schoenmakers, C.B. Lucasius and L. Buyden, *Chromatographia*, 36 (1993) 83.
54 L.R. Snyder, in Cs. Horváth (Editor), *High-Performance Liquid Chromatography, Advances and Perspectives, Vol. 1*, Academic Press, New York, 1980, p. 207.
55 V. Packova, K. Stulik and M. Prihoda, *J. Chromatogr.*, 442 (1988) 147.