Journal of Chromatography, 485 (1989) 433–451
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 21 983

COMPUTER-ASSISTED OPTIMIZATION WITH NEMROD SOFTWARE

G. MAZEROLLES, D. MATHIEU, R. PHAN-TAN-LUU and A. M. SIOUFFI*
Faculté des Sciences de Saint-Jerôme, Université Aix-Marseille III, F-13397 Marseille Cedex 13 (France)

SUMMARY

A computer-assisted optimization procedure for reversed-phase high-performance liquid chromatography which makes use of NEMROD software is described. The starting experimental domain is a tetrahedron, each vertex of which is a pure component. Truncations of the tetrahedron are performed either with available information or with a screening matrix. The result is either an irregular polyhedron which permits the selection of a design matrix or a symmetrical polyhedron. In the latter instance the influence of each solvent is easy to check and the experiments to be performed are less numerous than in the previous method. A new response function, I_c, derived from the information theory is described. Calculation of the I_c function is carried out in the selected domain and the maximum value is determined which permits the selection of a suitable mixture of the components of the mobile phase. Methoxyflavones and vanilla compounds are considered as examples.

INTRODUCTION

In liquid chromatography, finding the optimum solvent composition for the mobile phase is often a major task. Complete resolution of the solutes in a complex mixture is usually not possible with simple binary mixtures of solvents and selectivity "tuning" is achieved with multi-component solvents. Selection of appropriate proportions of these solvents is tedious and optimization is required. As alkyl-bonded silica supports are the most popular, advocated optimization schemes are mainly devoted to reversed-phase high-performance liquid chromatography (RP-HPLC).

According to Snyder[1], mixtures of four solvents (acetonitrile–methanol–tetrahydrofuran–water) in various proportions in most instances permit the widest selectivity range to be obtained. In a mixture experiment the response from a mixture of q components is a function of the proportions $X_1, X_2, \ldots, X_q$ of the components of the mixtures. The proportions X_i must satisfy the constraints $\Sigma X_i = 1$ and $0 \leqslant X_i \leqslant 1$. Selection of the X_i values requires first the determination of the experimental domain. This can be carried out with thin-layer chromatography or gradient runs or some selected isocratic HPLC experiments. With the domain the analyst must select a response function which describes his aims and takes into account all parameters and the interactions between them. The response function must take into account the priority goals, which are the analysis time and the quality of the separation between each pair of solutes.

0021-9673/89/$03.50

Many criteria for comparing the quality of chromatograms and response functions have been described and reviewed[2–5]. One strategy is to emphasize the resolution $R_s = 2(t_{r_j} - t_{r_i})/(\omega_i + \omega_j)$ between two adjacents eluites and another is to emphasize a discrimination factor as was proposed by Martin and Elfallah[6]. The search for the optimum of the response function within the experimental domain makes use of mixture experimental designs and is possible only with computer-assisted procedures. Many have been published and reviewed[2,7–9]. The Sentinel System derived from the pioneering work of Glajch *et al.*[10] appears to be the most sophisticated and the ORM (overlapping resolution map) method is widely advocated. In this mode water (the less eluting solvent) can be considered as a diluent whereas acetonitrile, methanol and tetrahydrofuran (THF) account for the proton donor, proton acceptor and dipole–dipole interactions. These conditions define a $(q-1)$ dimensional simplex and in the ORM procedure a triangle is constructed, the summits of which are mixtures of water and organic modifiers selected in such a way that the eluting strength of these three mixtures would be equivalent. In this paper we do not consider water as a diluent but as an actual solvent. In this mode we can first build a tetrahedron, each of the four vertices of which is a pure solvent, and it represents all possible combinations of the four solvents selected. In most instances this tetrahedron is far too large and we can perform truncations to select the experimental domain of interest, which can exhibit either a remaining symmetry or no symmetry at all. Inside this working volume the eluting strength may vary widely but the elution of all the solutes is achieved in an acceptable analysis time. The response function is optimized in the domain and the selected solvent mixture is checked. Two examples are given to demonstrate the feasibility of the method: the separation of polymethoxylated flavones and the separation of compounds from vanilla flavour. In both instances the number of solutes in the mixture is known. This is not always true when the analyst deals with a mixture to be separated, but in most published optimization procedures this is the case.

THEORY

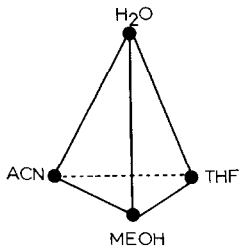
The screening matrix

As we selected optimization in RP-HPLC with the four solvents previously stated, the starting experimental domain is the tetrahedron displayed in Fig. 1. Two situations are possible:

(i) When retention data are available from the literature, truncations can be carried out (Fig. 2). If necessary the procedure of MacLean and Anderson[11] can be used. No experiment is needed to define the working volumes (this is the case for the flavone separation).

(ii) More generally information is scarce. Most retention data are given for binary mixtures[12]. Some preliminary experiments are therefore needed. A screening matrix according to the procedure of Snee and Marquardt[13] is constructed and reported in Table I. The purpose is to select a working domain in which the capacity ratio of the solutes remains in an acceptable range.

Thirteen mixtures of solvents were selected to ensure uniform coverage of the experimental domain. The meaning of the variables X_i is straightforward when considering the whole tetrahedron: lines 1–4 correspond to the use of a single solvent,



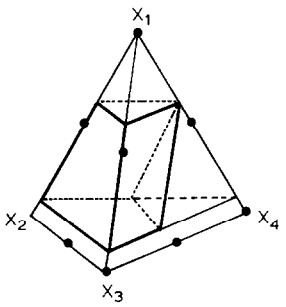


Fig. 1. The starting experimental domain. ACN = Acetonitrile; MEOH = methanol; THF = tetrahydrofuran.

Fig. 2. Truncations of the starting tetrahedron yielding an irregular polyhedron (methoxyflavone separation).

lines 5–8 correspond to mixtures with one solvent in a proportion of 62.5%, lines 9–12 correspond to mixtures of three solvents in equal proportions and the last line corresponds to equal proportions of the four solvents.

Chromatographic runs according to the selected X_is of the screening matrix are carried out and the capacity factor k'_j of every solute is recorded. The influence of a single solvent is determined by the observed variations in k' when its relative proportion in the solvent mixture is changed. A survey of the published applications of optimization procedures, particularly the ORM technique, indicated that in many instances the optimum is located in the close vicinity of one edge of the triangle. This may indicate that in those particular separations one solvent plays a minor role and is needed only for fine tuning of the selectivity.

As four components are involved, the simplest way to construct the screening matrix is to consider the pseudo-components displayed in Table II and to perform experiments.

TABLE I

THE SCREENING MATRIX

X_1, X_2, X_3 and X_4 are coded variables (for the meaning see text).

No.	X_1	X_2	X_3	X_4
1	1.000	0.000	0.000	0.000
2	0.000	1.000	0.000	0.000
3	0.000	0.000	1.000	0.000
4	0.000	0.000	0.000	1.000
5	0.625	0.125	0.125	0.125
6	0.125	0.625	0.125	0.125
7	0.125	0.125	0.625	0.125
8	0.125	0.125	0.125	0.625
9	0.000	0.333	0.333	0.333
10	0.333	0.000	0.333	0.333
11	0.333	0.333	0.000	0.333
12	0.333	0.333	0.333	0.333
13	0.250	0.250	0.250	0.250

TABLE II

THE 97:1:1:1 (v/v/v/v) PSEUDO-COMPONENTS

X_{H_2O}	X_{CH_3CN}	X_{CH_3OH}	X_{THF}
0.970	0.010	0.010	0.010
0.010	0.970	0.010	0.010
0.010	0.010	0.970	0.010
0.010	0.010	0.010	0.970

More precise information can be obtained in some particular instances by changing the proportions of the solvents of the pseudo-components. For example, with the vanilla compounds it is obvious from their structure and some previous data that they would not be retained with a pure organic solvent. A preliminary experiment with methanol–water (50:50, v/v) was carried out as it represents the middle of the edge. All the products were poorly retained and the separation could be tedious or impossible.

From this preliminary experiment, it was possible to discard solvent mixtures in which the proportion of organic modifier is higher than 50%. New X_i values are defined and the S' matrix is displayed in pure components in Table III and the corresponding experimental points in Fig. 3.

The influence of a single solvent i is determined by considering the experimental points from the vertex ($X_i = 1$) corresponding to this solvent to the centroid of the opposed face ($X_i = 0$). In this mode the observed data corresponding to (1,5,13,9), (2,6,13,10), (3,7,13,11) and (4,8,13,12) provide information on the influence of water, acetonitrile, methanol and tetrahydrofuran, respectively. For example, plots of the capacity factor of ethylvanillin are displayed in Fig. 4.

The above procedure permits the acceptable retention range to be selected. By

TABLE III

THE S' SCREENING MATRIX WITH NATURAL VARIABLES

No.	X_{H_2O}	X_{CH_3CN}	X_{CH_3OH}	X_{THF}
1	0.970	0.010	0.010	0.010
2	0.500	0.500	0.000	0.000
3	0.500	0.000	0.500	0.000
4	0.500	0.000	0.000	0.500
5	0.790	0.070	0.070	0.070
6	0.560	0.310	0.065	0.065
7	0.560	0.065	0.310	0.065
8	0.560	0.065	0.065	0.310
9	0.500	0.166	0.166	0.166
10	0.660	0.000	0.170	0.170
11	0.660	0.170	0.000	0.170
12	0.660	0.170	0.170	0.000
13	0.620	0.126	0.126	0.126

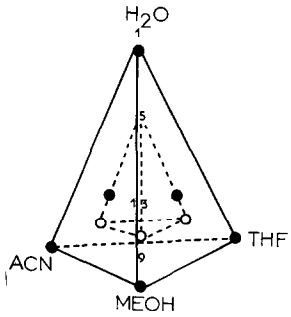


Fig. 3. Reduction of the tetrahedron from experiments corresponding to lines 2, 3 and 4 of the screening matrix and new tetrahedron in which the screening is performed (dashed lines).

overlapping these plots for each pair of solutes and for a selected solvent, we can check the selectivity $\alpha = k'_j/k'_i$ and hence the separation. It must be pointed out that four points of the experimental domain are involved and uniformly distributed. The general trend of the contribution of one solvent to different possible separations is evidenced but a fine prediction cannot be achieved. When two parallel lines are observed with one pair of solutes it means that this particular solvent does not affect the separation. It must be remembered that two parallel lines are not usual but rather an exception.

The response function

When a complex sample with n solutes is chromatographed, the chromatogram exhibits k_1 singlets, k_2 doublets and k_u u-plets are observed with $\sum_p pk_p = n$, where p is the number of solutes present in one given multiplet, varying from 1 to u.

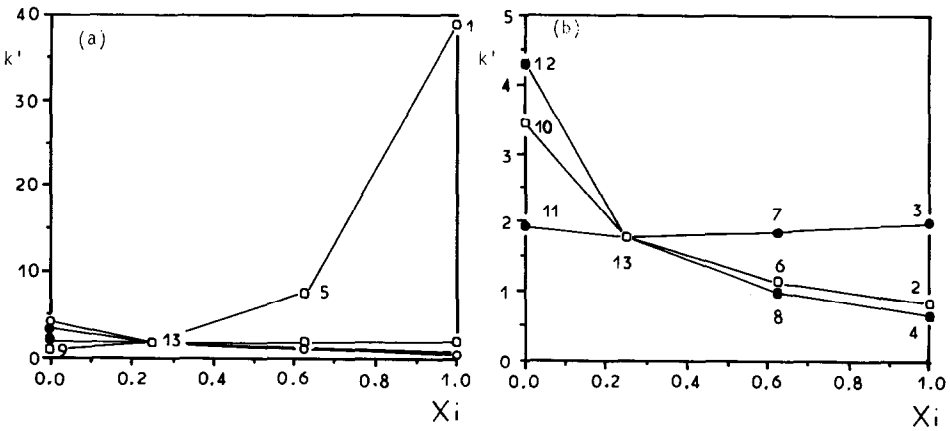


Fig. 4. Plots of capacity factor k' *versus* proportion of solvent i (solute, ethylvanillin; for chromatographic conditions, see text), (a) All solvents are considered; □ = water; ● = acetonitrile; ■ = methanol; ○ = THF. (b) Organic modifiers only are considered: □ = acetonitrile; ● = methanol; ■ = THF. The numbers on the different lines correspond to lines of the screening matrix.

From the information theory and the early work of Massart *et al.*[14], we can define the I_c function as

$$I_c = \sum_p k_p p/n \cdot \log_2(n/p)$$

where k_p is the number of p-plets in the chromatogram.

It is obvious that a solute is clearly identified when the observed resolution with neighbouring peaks is higher than 1.25. The amount of information resulting from the identification of a single compound is given by

$$I_1 = \log_2(n/1) = \log_2 n$$

The frequency of singlet appearance is given by $f_1 = k_1/n$, where f_1 is a measure of the probability for a single solute of the sample appearing as a singlet in the chromatogram.

The contribution of the k_1 singlets to the amount of overall information from the whole chromatogram is

$$I_s = (k_1/n)I_1 = (k_1/n)\log_2 n$$

By extrapolation it can be argued that for $2k_2$ compounds the number of possible identities shifts from n to 2. The frequency of doublet appearance is $f_2 = 2k_2/n$ and f_2 is the probability of one solute appearing as a doublet. The contribution of the k_2 doublets to the overall information is given by

$$I_d = (2k_2/n)\log_2(n/2)$$

The same reasoning can be carried out with the k_p p-plets in the chromatogram.

Consider a chromatogram of eight solutes exhibiting four singlets and 2 doublets; the value of the I_c function is

$$I_c = 4 \cdot 1/8 \log_2(8) + 2 \cdot 2/8 \cdot \log_2(4) = 1.5 + 1 = 2.5$$

The maximum value of I_c is obtained in a chromatogram that exhibits only singlets. In this instance $I_c = (n/n)\log_2 n = \log_2 n$. In the same way, the minimum value of I_c is obtained when all peaks are eluted at the same retention time and $I_c = \log_2(n/n) = \log_2(1) = 0$.

When the analyst knows the number of compounds in the sample (n), I_c can be calculated. The drawback of this response function is that I_c is not known when n is not available. As was stated above, in most instances the optimization procedures deal with a known number of solutes. This is a formidable task to optimize when the number of solutes is unknown and under the best conditions the analyst may miss some compounds.

Optimization: The NEMROD software[a]

The NEMROD software is used in the construction and treatment of matrices and is a very powerful tool in the methodology of experimental design. The goal of optimization is to select a set of experiments that will provide the best information with the highest accuracy. The set of experiments to carry out is generated with the NEMROD software. Different matrices of experiments are proposed: screening matrices (Hadamard, Plackett and Burmann), factorial matrices (2^k, 3^k, 2^{k-1}), matrices of response surface (Doehlert, Box-Benken, etc.), matrices of mixtures (Scheffe type) or any type of standard matrices[14]. From these matrices, a screening of the factors, a study of the interactions between these factors and a knowledge of the experimental domain can be carried out.

In some instances none of the above matrices is convenient and NEMROD permits the construction of a different matrix making use of different mathematical algorithms (exchange, uniform distribution). Moreover, some other constraints, such as time, methods or cost of experiment, can be taken into account. The software permits a mathematical model to be postulated to depict the phenomenon to study: polynomial model, synergistic model, mixture model (Scheffe, Draper–St. John) or any particular model. From the experimental data obtained, the software calculates the coefficients of the model by a multilinear regression. Owing to the mathematical model, the analyst can perform a screening among multifactors, determine the influence of selected factors, calculate the interactions between the different factors and calculate the responses in every part of the experimental domain. About 30 programs written in different languages (PASCAL, FORTRAN, ASSEMBLER) compose the NEMROD software.

When no symmetry in the experimental domain is observed, the use of Scheffe's experimental design[15] is not possible. As was described previously[16], a design matrix ξ is constructed from the polyhedron by considering the vertices, the overall centroid, the centroids of edges, faces and hyperfaces.

The response function in the different mixtures of solvents is an mth polynomial. For the sake of simplicity, a second-order polynomial is sufficient in most instances:

$$\eta = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{14} X_1 X_4 + \beta_{23} X_2 X_3 + \\ + \beta_{24} X_2 X_4 + \beta_{34} X_3 X_4$$

where β_i is the weight of pure i and β_{ij} is the quadratic coefficient of the binary synergism.

Calculation of the coefficients is carried out by the least-squares method using

$$\boldsymbol{B} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$$

in which $\boldsymbol{B}$ is the vector of the estimates of coefficients β, $(\boldsymbol{X}'\boldsymbol{X})$ is the information matrix and $(\boldsymbol{X}'\boldsymbol{X})^{-1}$, the inverse of $(\boldsymbol{X}'\boldsymbol{X})$, is the dispersion matrix.

We can select those experiments which give the highest information taking into account the following criteria: (i) maximum value of the determinant of the

[a] Available on request to Prof. Phan-tan-Luu.

information matrix; (ii) minimum value of the variance function within the whole domain which leads to a good quality of the estimate in the whole domain; and (iii) minimum set of experiments.

In this respect the Nemrod exchange algorithm permits the determination of the best compromise between the estimation of the response function and the set of experiments. The Nemrod exchange algorithm works according to the Fedorov method[17]. We start the procedure with a given number of candidate points n' obtained by considering the points of the design matrix ξ. This number is much larger than the number of the p coefficients of the Y polynomial. In order to select those experiments which yield the best information, the moment matrix $M = (X'X)/z$ is defined, where z is the number of experiments of the X matrix. From the minimum set of experiments as defined by the number of the p coefficients of the mathematical model (ten for the second-order polynomial above) to n', the Nemrod exchange algorithm calculates the highest value of the determinant of $(X'X)$ for every set of experiments. An histogram of the M determinant versus z yields the number and the nature of experiments to perform. Selected experiments give the highest values of M.

EXPERIMENTAL

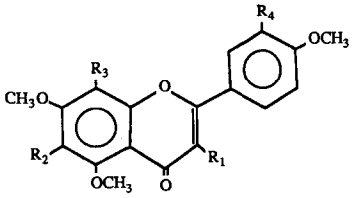
Chemicals

Structures of the methoxyflavones are given in Table IV. There were gifts from the Laboratoire de Phytochimie of the University of Aix-Marseille and were isolated from citrus or synthesized by the procedure of Gaydou and Bianchini[18].

Anisaldehyde (An. Ald.), protocatechuic aldehyde (Pr. Ald.), *p*-hydroxybenzaldehyde (P.H.B. Ald.), veratraldehyde (Ver. Ald.), vanillin (Van.), protocatechuic acid (Pr. Ac.), *p*-hydroxybenzoic acid (PHB. Ac.), vanillic acid (V. Ac.), vanillyl alcohol (V.

TABLE IV

STRUCTURES OF THE POLYMETHOXYLATED FLAVONES INVESTIGATED



No.	*Name*	R_1	R_2	R_3	R_4
1		H	H	OCH_3	OCH_3
2		OCH_3	H	OCH_3	OCH_3
3	Sinensetin	H	OCH_3	H	OCH_3
4		OCH_3	OCH_3	H	OCH_3
5		OCH_3	OCH_3	H	H
6	Tangeritine	H	OCH_3	OCH_3	H
7	Nobiletin	H	OCH_3	OCH_3	OCH_3
8	Heptamethoxyflavone	OCH_3	OCH_3	OCH_3	OCH_3

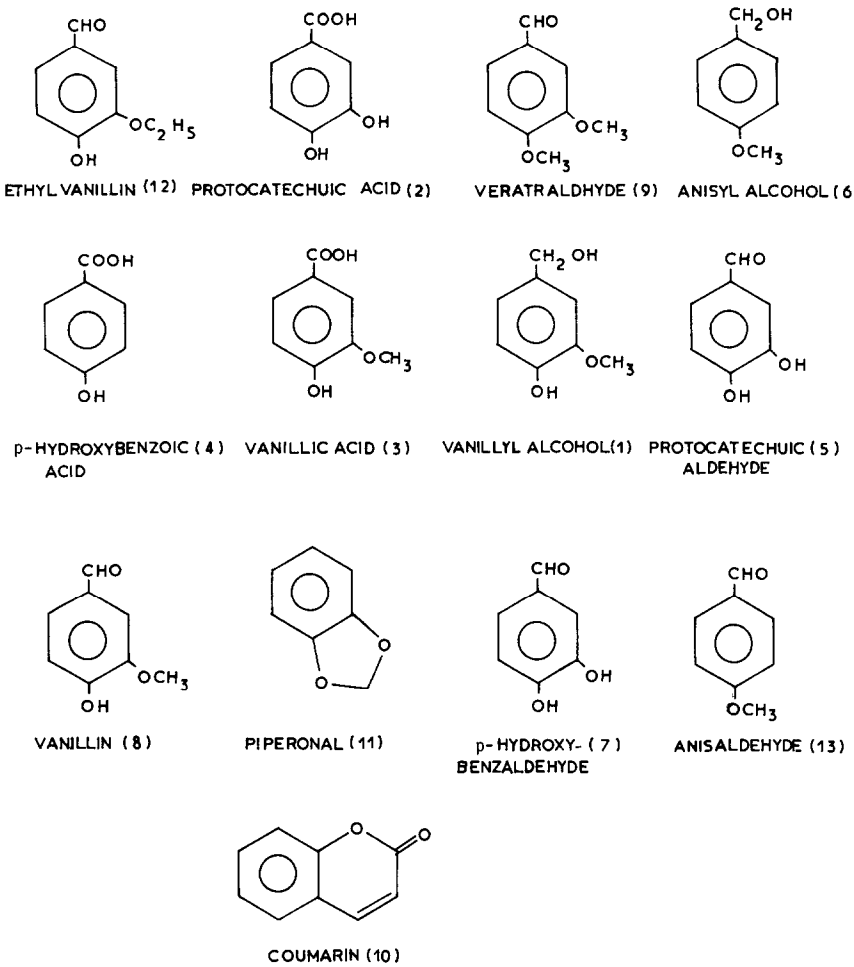


Fig. 5. Structures of the vanilla compounds.

Al.), anisyl alcohol (An. Al.), ethylvanillin (Etvan.), coumarin (Cou.) and piperonal (Pip.) were purchased from Fluka (Buchs, Switzerland). The structures are shown in Fig. 5.

Solvents of LiChrosolv grade were obtained from Merck (Darmstadt, F.R.G.). Water was distilled over potassium permanganate and purified by percolation through a Lobar column (Merck) packed with RP-18 particles.

Chromatography

Experiments were performed with a Merck Hitachi 655A chromatograph equipped with a Rheodyne 7125 sample loop (20 μl) and a fixed-wavelength detector (254 nm). Columns were either 250 × 4 mm I.D. for methoxyflavones or 125 × 4 mm I.D. packed with 5-μm LiChrosorb RP-18 particles for vanilla compounds. The flow-rate was 1 ml/min.

RESULTS AND DISCUSSION

Separation of methoxyflavones

Truncation of the starting tetrahedron was easy to perform as some previous experiments[19] had demonstrated that beyond 60% water in the eluent mixture the observed k' values were too high. In this instance the design matrix has 25 lines. As the response function is a second-order polynomial with ten coefficients, a minimum set of ten experiments must be performed. Thirteen were determined by the Nemrod exchange algorithm. Sections of the experimental domain are difficult to perform and to facilitate the visualization of the response, the experimental matrix is transformed in such a way that any value of the last column is zero. In this mode a cube overlaps the tetrahedron. Sections of the cube are performed assuming one variable is kept constant and the experimental domain is easily visualized in Fig. 6 (for details, see ref. 16).

In the section corresponding to the experimental domain, isoresponse curves are

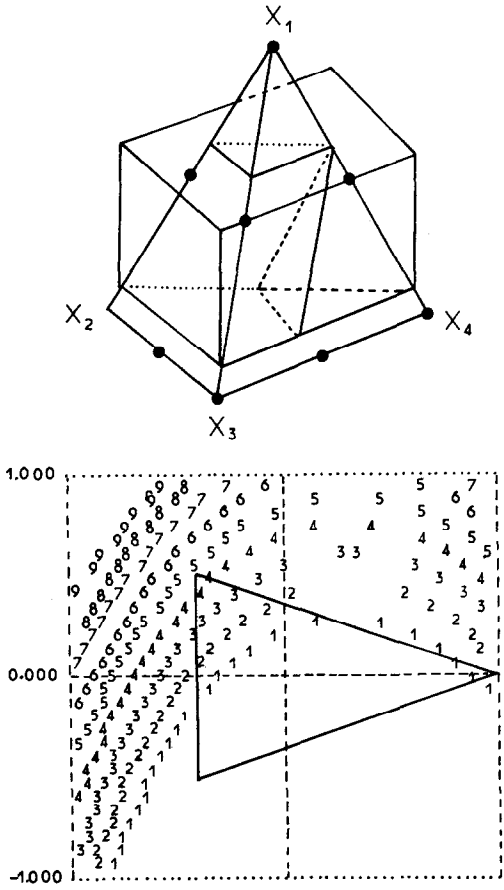


Fig. 6. Overlapping of the polyhedron with a cube of "similar" dimensions and representation of one section with isoresponse curve.

TABLE V

OBSERVED I_c VALUES FOR METHOXYFLAVONES IN TABLE III (IN ACTUAL VARIABLES)

Pt Nb	X_{H_2O}	X_{THF}	X_{CH_3CN}	X_{CH_3OH}	I_c
1	0.100	0.000	0.000	0.900	0.536
2	0.100	0.400	0.000	0.500	0.000
4	0.600	0.400	0.000	0.000	1.904
5	0.100	0.000	0.900	0.000	0.000
6	0.100	0.400	0.500	0.000	0.000
7	0.600	0.000	0.400	0.000	2.504
8	0.100	0.200	0.000	0.700	0.000
9	0.350	0.000	0.000	0.650	1.808
10	0.100	0.000	0.450	0.450	0.000
11	0.350	0.400	0.000	0.250	0.000
13	0.600	0.200	0.000	0.200	2.504
18	0.350	0.000	0.650	0.000	1.436
24	0.350	0.200	0.450	0.000	1.119

drawn for every pair of solutes and a procedure similar to the ORM method of Glajch *et al.*[10] is performed in order to select the optimum value.

With this method the optimum composition is THF–acetonitrile–methanol–water (0.15:0.05:0.22:0.58). In this area all solutes are well separated and the observed resolutions are higher than 1. In this procedure resolution was the criterion for the response function. We compared these previous results with the I_c function described above. In this mode, calculated I_c values from the 13 experimental runs defined by the matrix are calculated and are reported in Table V. The I_c value corresponding to the previous optimum response function is then calculated ($I_c = 2.75$) and corresponds to a chromatogram exhibiting six singlets and one doublet. This is evidenced in the chromatogram in Fig. 7 as peaks 3 and 4 are not completely resolved ($R_{3-4} = 1.1$). A higher methanol content will provide higher I_c values but will yield much higher k'

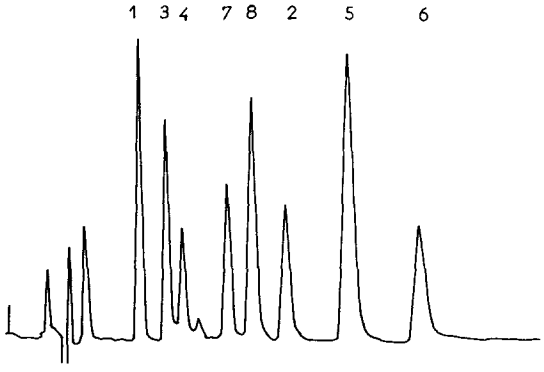


Fig. 7. Separation of the methoxyflavones with RP-HPLC column (250 × 4 mm I.D.) Hibar packed with RP-18 (5 μm) particles. Mobile phase, THF–CH_3CN–CH_3OH–H_2O (15:5:22:58, v/v/v); flow-rate, 1 ml/min; ΔP = 150 bar; detection, UV (254 nm). Numbers correspond to structures in Table IV.

TABLE VI

OBSERVED CAPACITY FACTORS OF THE COMPOUNDS IN TABLE IV (Pt Nb CORRESPONDS TO LINE OF THE SCREENING MATRIX)

For identification of the solutes, see Experimental.

Pt Nb	*Pr. Ac.*	*V. Al.*	*PHB. Ac*	*V. Ac.*	*Pr. Ald.*	*PHB. Al.*	*An. Al.*	*Van.*	*Pip.*	*Ver. Ald.*	*Cou.*	*Etvan.*	*An. Ald.*
1	3.00	4.00	6.00	6.50	8.75	13.75	15.25	17.75	27.75	30.00	31.25	39.00	41.62
2	0.06	0.20	0.40	0.06	0.46	0.60	0.60	0.53	1.33	1.26	1.26	0.86	1.66
3	0.27	0.30	0.73	0.65	0.45	1.00	1.00	1.06	3.26	2.01	3.00	2.00	3.86
4	0.18	0.38	0.60	0.38	0.43	0.58	0.48	0.53	0.89	0.69	0.79	0.69	0.94
5	1.33	1.13	2.53	1.90	2.88	5.06	4.77	4.33	9.66	5.46	8.62	7.73	8.80
6	0.36	0.19	0.40	0.24	0.43	0.85	0.83	0.64	1.91	1.21	1.35	1.16	1.97
7	0.24	0.29	0.53	0.45	0.66	1.21	1.30	1.10	2.64	1.26	1.96	1.86	2.80
8	0.40	0.60	0.66	0.60	0.66	0.86	0.66	0.86	1.33	1.80	1.06	1.00	1.26
9	0.20	0.20	0.60	0.33	0.46	0.73	0.60	0.66	1.26	0.86	1.06	0.93	1.46
10	0.98	0.58	1.80	1.18	1.83	2.93	1.81	2.03	4.38	1.77	2.81	3.44	4.80
11	0.46	0.60	0.86	0.53	0.93	1.33	1.26	1.13	2.60	1.40	2.00	1.93	2.73
12	0.33	0.60	0.60	0.60	1.06	2.86	2.73	2.20	5.73	7.06	5.80	4.29	7.40
13	0.36	0.37	0.69	0.46	0.80	1.33	1.20	1.04	2.68	1.33	2.00	1.78	2.86

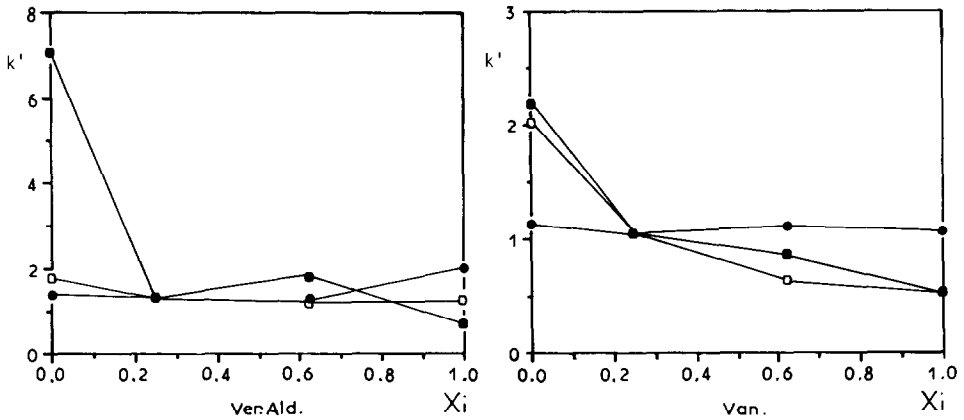


Fig. 8. Plots of k' *versus* X_i for two selected solutes, veratraldehyde (Ver. Ald.) and vanillin (Van.). These plots are similar to those in Fig. 4b. □ = Acetonitrile; ● = methanol; ■ = THF.

values beyond the selected analysis time. It may be pointed out that the acetonitrile content is very low.

Vanilla compounds

As less information is available on the retention of these solutes, to perform the same type of truncations as above we must use the screening matrix with the solvent mixtures given in Table III. Experiments were carried out and the capacity factors were recorded (Table VI). Plots similar to those in Fig. 4 were drawn for every solute. As could be guessed, the water content has a major influence and it is necessary to work with a minimum of 30% of water to obtain retention. From the plots it can be demonstrated that some solutes, such as veratraldehyde, are very sensitive to the THF content and others, such as 4-hydroxybenzoic acid, to the acetonitrile content. Surprisingly, the slopes of the lines observed when focusing the methanol influence were only slightly different. This means that methanol plays only a minor role (an example is given in Fig. 8). This result is consistent with the previous work on an ORM optimization technique for seven of these compounds[20]. It was observed that the

TABLE VII

SCHEFFE'S MATRIX IN THE NEWLY DETERMINED EXPERIMENTAL DOMAIN

X_i values are coded variables.

No.	X_1	X_2	X_3
1	1.000	0.000	0.000
2	0.000	1.000	0.000
3	0.000	0.000	1.000
4	0.500	0.500	0.000
5	0.500	0.000	0.500
6	0.000	0.500	0.500
7	0.333	0.333	0.333

TABLE VIII

OBSERVED CAPACITY FACTORS FOR THE VANILLA SOLUTES IN EVERY POINT CORRESPONDING TO TABLE VII

Pt Nb	*Pr. Ac.*	*V. Al.*	*PHB. Ac*	*V. Ac.*	*Pr. Ald.*	*PHB. Al.*	*An. Al.*	*Van.*	*Pip.*	*Ver. Ald.*	*Cou.*	*Etvan.*	*An. Ald.*
1	2.20	3.20	4.10	4.30	6.45	10.20	11.80	13.93	22.73	24.86	25.93	31.80	33.76
2	1.37	0.58	2.15	1.42	2.15	3.16	1.79	2.15	4.58	1.63	3.05	3.63	4.68
3	0.10	0.61	0.16	0.40	0.88	1.80	2.16	1.88	5.60	3.70	5.27	3.70	6.70
4	2.10	1.00	3.80	2.60	4.05	6.45	3.70	5.00	9.25	4.10	7.00	9.50	10.70
5	0.20	1.10	0.75	0.75	2.20	3.90	5.00	4.70	13.06	9.75	13.00	10.86	17.00
6	0.30	0.45	0.50	0.40	1.30	2.05	1.75	1.85	4.00	2.00	3.20	3.25	4.30
7	0.50	0.75	1.05	0.75	2.00	3.40	2.80	3.00	6.40	3.20	5.45	5.60	7.50

optimum was located very close to the line connecting water–acetonitrile vertex to water–THF vertex in the triangle. Suppressing of methanol will certainly affect the resolution and it should not be done in a general scheme. However, in this particular instance we can suppress methanol by considering that its influence is very minor and that a ternary mixture is simpler to handle than a quaternary system. Mixtures of water, acetonitrile and THF are sufficient to obtain large selectivity changes and the domain of interest is thus a triangle. To retain symmetry and in order for the capacity factors to remain in acceptable range the new vertices were defined as follows:

X_1 = water–acetonitrile–THF (0.96:0.02:0.02)
X_2 = water–acetonitrile–THF (0.73:0.01:0.26)
X_3 = water–acetonitrile–THF (0.73:0.26:0.01)

A mixture design according to Scheffe is constructed in the domain of interest (Table VII). The response function is a simplified cubic model:

$$Y = b_1X_1 + b_2X_2 + b_3X_3 + b_{12}X_1X_2 + b_{13}X_1X_3 + b_{23}X_2X_3 + b_{123}X_1X_2X_3$$

The seven experiments corresponding to this matrix are carried out and the recorded k' values are given in Table VIII. From the data obtained it is observed that many peak

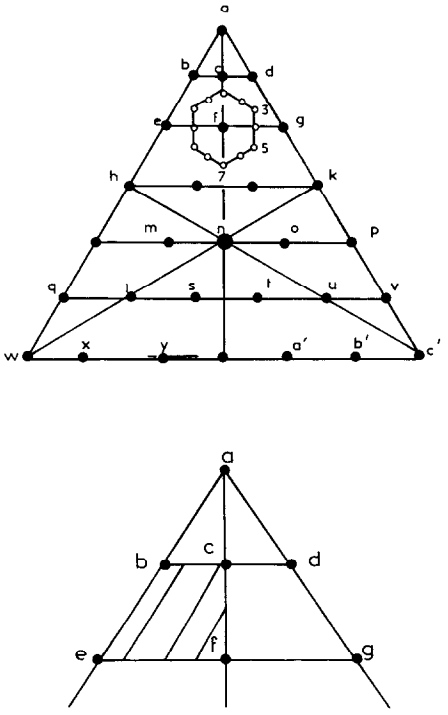


Fig. 9. Vanilla compounds. (a) Distribution of the experimental points in the working domain. a, w and c' correspond to lines 1, 2 and 3 in Table VII. (b) Magnification of the selected area in which twelve peaks are clearly separated and identified in the chromatogram.

crossovers occur. Twenty peak crossovers were determined from the observed order of retention with the mixture corresponding to the water vertex (point a in Fig. 9). This demonstrates that the screening matrix allowed the selection of those solvents which provide maximum selectivity.

Owing to this very large number of peak crossovers, the modelling of the I_c function or any global response function is almost impossible to achieve as many discontinuities occur. In the same way, an attempt to use the ORM method is time consuming. A straightforward procedure is to perform a model of the k' variations within the experimental domain. The knowledge of $\log(k') = f(X_1, X_2, X_3)$ permits the order of elution and the resolution between each pair of solutes to be determined in every part of the experimental domain. Resolution (R_s) is calculated through

$$R_s = (K'_i - K'_j)/(K'_i + K'_j + 2) \cdot \sqrt{N}/2$$

In this equation the plate count is taken as the lowest value obtained in the preliminary experiments. For the sake of simplicity, it was considered that N is constant. This is not true but as we used the lowest value, we can consider that the calculation is performed in the most unfavourable conditions. It may be stated that $R_s = 1$ would be considered as correct for separation and is used as the threshold for I_c. In other words, when two peaks are separated with $R_s = 1$ there are considered as two singlets; when $R_s < 1$ the peaks are not resolved and they are considered as a doublet. The maximum value of the I_c function is $I_{c\,max} = 13 \cdot 1/13 \cdot \log_2(13) = 3.692$. This corresponds to a chromatogram in which the observed resolutions are all higher than 1. From the estimates of the resolution R_s, the I_c function is calculated. One example is given in Table 9 for point e in Fig. 9.

The I_c function is calculated for 34 evenly distributed points, which means that 41 values of I_c are either experimentally determined or calculated. In the area where the

TABLE IX

CALCULATION OF THE I_c FUNCTION FOR POINT e OF FIG. 9

Mobile phase: water–THF (2:1).

Compound	*Log* k'_{est}	k'_{est}	R_{calc}	*Peak*	*Partial* I_c
V. Al.	0.3191	1.38		1	0.284
Pr. Ac.	0.8657	2.38	5.50	1	0.284
V. Ac.	1.2049	3.34	3.93	1	0.284
PHB. Ac	1.4109	4.10	2.56	1	0.284
Pr. Ald.	1.5751	4.83	2.12	[a]	0.000
An. Al.	1.6468	5.19	0.94	2	0.415
Ver. Al.	1.9141	6.78	3.60	[a]	0.000
Van.	1.9305	6.89	0.23	2	0.415
PHB. Ald.	2.0452	7.73	1.59	1	0.284
Cou.	2.3293	10.27	4.02	1	0.284
Pip.	2.5100	12.30	2.62	1	0.284
Ethvan.	2.6267	13.83	1.71	1	0.284
An. Ald.	2.7175	15.14	1.34	1	0.284
					I_c 3.386

[a] Not separated from the next solute.

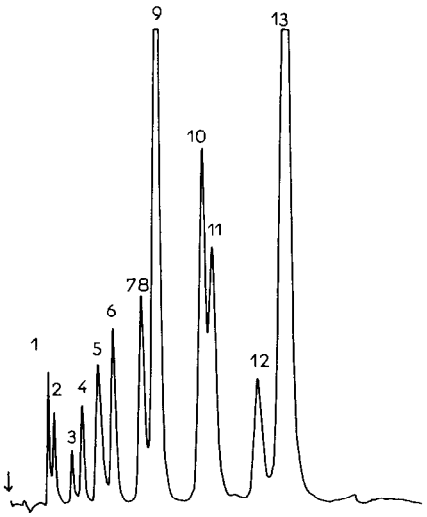


Fig. 10. Separation of vanilla compounds by RP-HPLC. Column, 125 × 4 mm I.D. Hibar packed with RP-18 (5 μm) particles. Mobile phase: THF–CH_3CN–H_2O (6:2:92, v/v/v); flow-rate, 1 ml/min; Δp = 98 bar; detection, UV (254 nm). Numbers correspond to structures in Fig. 5.

highest values of I_c are obtained a narrow mesh grid is performed. From the results it is concluded that complete separation of the thirteen compounds would not be achieved as the highest value of I_c is 3.539 (point b in Fig. 9). On the chromatogram displayed in Fig. 10, twelve peaks are observed. Calculations of the resolution between vanillin and 4-hydroxybenzaldehyde yields 0.40, which means that a singlet would be observed in every instance. An additional run with a pure standard demonstrated that the difference in k' is small (8.75–9.00). A more efficient column would achieve the complete separation. The critical point of the above procedure is that the void volume V_0 and as a consequence t_0 is considered to be constant whatever the proportions of the components of the eluent. To minimize the errors we determined t_0 with injection of pure water. In the same way, it may be argued that k' may vary with the column lifetime. We observed a slight decrease in k' with time but we ensured that the selectivity was kept constant.

To improve the estimated models, three chromatograms corresponding to points f, r and u in Fig. 9 were recorded. From the data, models similar to those previously utilized were constructed and the coefficients calculated through a least-squares linear regression with the ten experimental points (the seven previously used and the three from the last runs). The coefficients remained unchanged, which means that the previous model was correct (Table X). The I_c values obtained are identical with those previously obtained. We did not find any experimental conditions for separating the thirteen solutes with this 12.5-cm column, but we could determine a zone in which twelve peaks are well separated and identified. A magnification of this zone is displayed in Fig. 9b.

TABLE X

COMPARISON OF THE CALCULATED AND OBSERVED I_c FUNCTIONS FOR POINT b IN FIG. 9

Retention order	k'	R_{exp}	*Calc. retention order*	*Log* k'_{est}	k'_{est}	R_{calc}	$R_{calc} - R_{exp}$
V. Al.	2.10		V. Al.	0.7066	2.03		
		0.99				1.77	0.78
Pr. Ac.	2.30		Pr. Ac.	0.8705	2.39		
		5.86				5.54	−0.32
V. Ac.	3.80		V. Ac.	1.344	3.83		
		1.11				1.11	0.00
PHB. Ac.	4.15		PHB. Ac.	1.4293	4.18		
		3.86				3.91	0.05
Pr. Ald.	5.58		Pr. Ald.	1.7302	5.64		
		1.07				1.37	0.30
An. Al.	6.04		An. Al.	1.8302	6.24		
		5.10				5.08	−0.02
PHB. Ald.	8.75		PHB. Ald.	2.1979	9.01		
		0.40				1.05	0.65
Van.	9.00		Van.	2.2719	9.70		
		1.07				1.14	0.07
Ver. Ald.	9.70		Ver. Ald.	2.3512	10.50		
		6.15				6.00	−0.15
Cou.	14.87		Cou.	2.7656	15.89		
		0.79				0.82	0.03
Pip.	15.68		Pip.	2.8206	16.79		
		3.23				3.12	−0.11
Ethvan.	19.48		Ethvan.	3.0293	20.68		
		1.44				1.23	−0.21
An. Ald.	21.43		An. Ald.	3.1104	22.43		

CONCLUSION

Optimization of a quaternary solvent mobile phase for reversed-phase chromatography can be achieved within a wide experimental domain with the help of the I_c response function and the Nemrod exchange algorithm. Combination of the I_c function and Nemrod exchange algorithm. Combination of the I_c function and Nemrod software permit the optimization of the mobile phase within a wide experimental domain. The influence of one solvent may be emphasized by a screening matrix which may be used to obtain a further insight into the retention mechanism. When many peak crossovers occur, a response function based on the criteria of resolution is unsuitable and the only way to overcome this difficulty is to model log k' in order to optimize the selectivity.

REFERENCES

1 L. R. Snyder, *J. Chromatogr. Sci.*, 16 (1978) 223.
2 P. J. Schoenmakers, *Optimization of Chromatographic Selectivity (Journal of Chromatography Library*, Vol. 35), Elsevier, Amsterdam, 1987.
3 H. J. G. Debets, *J. Liq. Chromatogr.*, 8 (1985) 2725.

4 P. J. Schoenmakers, *J. Liq. Chromatogr.*, 10 (1987) 1865.
5 J. C. Berridge, *Anal. Chim. Acta*, 191 (1986) 243.
6 M. Martin and M. R. Elfallah, *Analusis*, 16 (1988) 241.
7 M. A. Quarry, R. L. Grob and L. R. Snyder, *Anal. Chem.*, 58 (1986) 907.
8 C. E. Goewie, *J. Liq. Chromatogr.*, 9 (1986) 907.
9 J. L. Glajch, J. J. Kirkland and J. M. Minor, *J. Liq. Chromatogr.*, 10 (1987) 1727.
10 J. L. Glajch, J. J. Kirkland, R. M. Squire and J. M. Minor, *J. Chromatogr.*, 199 (1980) 57.
11 R. A. MacLean and V. Anderson, *Technometrics*, 8 (1966) 447.
12 J. P. Bounine and G. Guiochon, *Analusis*, 12 (1984) 175.
13 R. P. Snee and D. W. Marquardt, *Technometrics*, 18 (1976) 19.
14 D. L. Massart, A. Dijkstra and L. Kaufman, *Evaluation and Optimisation of Laboratory Methods and Analytical Procedures*, Elsevier, Amsterdam, 1978, pp. 166 and 243.
15 H. Scheffe, *J. R. Stat. Soc.*, B20 (1958) 344.
16 J. P. Bianchini, E. M. Gaydou, A. M. Siouffi, G. Mazerolles, D. Mathieu and R. Phan-tan-Luu, *Chromatographia*, 23 (1987) 15.
17 V. V. Fedorov, in W. J. Klimko and E. M. Klimko (Editors), *Theory of Optimal Experiments*, Academic Press, New York, 1972.
18 E. M. Gaydou and J. P. Bianchini, *Bull. Soc. Chim. Fr.*, 11 (1978) 43.
19 S. Ramanarivo, *Ph.D. Thesis*, University Aix-Marseille III, Marseille, 1983.
20 A. M. Siouffi, *Chim. Mag.*, 7 (1986) 120.