Journal of Chromatography, 550 (1991) 27–37
Elsevier Science Publishers B.V., Amsterdam

CHROMSYMP. 2329

Fully automatic high-performance liquid chromatographic optimization

N. M. DJORDJEVIC*, F. ERNI and B. SCHREIBER
Sandoz Pharma Ltd., Analytical Research and Development, Basle (Switzerland)
E. P. LANKMAYR and W. WEGSCHEIDER
Graz University of Technology, Graz (Austria)
and
L. JAUFMANN
Kontron, Munich (Germany)

ABSTRACT

A good optimization routine should correctly find the best chromatographic separation conditions for mixtures of known or unknown constituents. The chromatographer must define the number of the parameters to be optimized and their ranges. However, the more parameters to be optimized and the more they interact, the more difficult and time-consuming the optimization procedure will be. A system capable of performing fully automated optimization of mobile phase selectivity in reversed-phase liquid chromatography was built. The optimization routine searches for the best conditions (trying to maximize a chromatographic response function) and for the points, inside defined experimental borders, where the least available experimental information is available. By conducting the experiment under the predicted optimum conditions and an additional experiment under conditions corresponding to the least density of information, the system was forced not to search for a local maximum, but to approach the global optimum. Peak tracking, an important part of any optimization process in high-performance liquid chromatography, was an integral part of the optimization software and was based on fuzzy theory. This implementation of an on-line identification of the sample components made a fully automated optimization of the mobile phase composition possible. Once a suitable separation had been achieved, it was necessary to validate the procedure, special attention being focused on robustness. The robustness test appraises the outcome of small variations in method conditions on the analytical performance. An important feature of this robustness analysis was the three-dimensional representation of the data as the hypersurface which helps to relate robustness to elution characteristics.

INTRODUCTION

The use of computer-aided procedures for the optimization of separation selectivity in reversed-phase high-performance liquid chromatography (HPLC) has been extensively studied during the last decade, and various approaches are available to rationalize development and optimization. Excellent overviews on optimization techniques in chromatography can be found in the literature [1–3]. Owing to the wide variety of separation principles accessible for HPLC separations, systematic method

0021-9673/91/$03.50

development may help greatly in finding suitable experimental conditions for a given separation task. The lack of mechanistic models suitable for general application makes empirical optimization schemes the method of choice for unrestricted optimization in HPLC.

Systematic method development resulting in fully automated instrumentation is only useful if all of the variables which permit separation selectivities and retention times to be tuned effectively can be considered. Thus, a sequential approach based on a moving least-squares interpolation between experimentally obtained retention data has been proposed [4]. The moving least-squares algorithm is employed for a non-linear interpolation between experimentally obtained retention data, supplying retention curves for all sample components. The resulting data allow prediction of the separation for any intermediate experimental conditions. The resulting optimization program (OPTIM) calculates systematically the optimization hypersurface according to a preselected optimization criterion which reflects the quality of the separation.

In this study a chromatographic response function (CRF) derived from the peak separation factor [5] and generalized for multi-component analysis [6] was used with (CRF_1) and without (CRF_2) normalization on total analysis time.

$$CRF_1 = 1/t[\prod f/(g + 2n)] \quad (1)$$

$$CRF_2 = \prod f/(g + 2n) \quad (2)$$

where f and g are the separation factors according to Kaiser [5] and Wegscheider *et al.* [6], n is a baseline noise and t is analysis time, *i.e.* duration of the actual chromatogram.

The incorporation of peak separation, noise and analysis time information allows one to account for basic performance characteristics of analytical separation procedures, such as accuracy, precision, speed and ultimately also cost of analysis. Another advantage of this function is its adaptability to the actual chromatographic situation. It considers automatically the status of peak separation in the presence of peak asymmetry and at differing peak intensity ratios. The OPTIM program provides output information containing the proposed best experimental conditions and also gives the point in space which has the lowest density of experimental data. Consequently, information on the retention behavior of all analytes is accumulated in a stepwise manner to solve the apparent separation problem or to quit according to a predefined stop criterion. Owing to the sequential concept of this optimization strategy the total experimental effort needed for method development depends largely on the complexity of the separation problem itself. Simple separation tasks may call for only a few runs, whereas more complex problems may require a larger number of experiments.

For automated optimization it is necessary to keep track of the identity of the eluted signals. Typical approaches for peak identification include the separate injection of standard components and/or multi-wavelength detection with factor analysis [7]. The mere use of standard components makes the process extremely slow and can, of course, not be applied if sample components of unknown identity are analyzed. Overlapping peaks have been successfully deconvoluted using diode array detection. The use of strict statistical models for data analysis may, however, not be

adequate for fully automated peak recognition at greatly changed mobile phase compositions.

To provide a generally applicable procedure, a peak-tracking routine based on fuzzy theory has been developed. Fuzzy theory is applied sucessfully when uncertain data are to be processed for analytical reasoning, as is the case with peak data obtained during method development procedures. The incorporation of this theory allows one to account for imprecise data originating from altered experimental conditions and to assign the membership of peak data to a sample component by means of computed membership function values. This routine works with either single- or multi-channel (*e.g.* diode array) detectors and can also be used with unknown compounds, even when peak overlap occurs during the optimization runs [8,9]. Recognition of peaks is based on comparison of peak areas and elution order of the signals from a reference and a trial run. As the number of signals in the chromatograms can alter during progress of the optimization, the chromatogram showing the maximum number of signals is always chosen as a reference. In the case of peak overlap, each potentially overlapped peak is compared with all linear combinations of the reference peaks not recognized as single peaks. Fuzzy comparison results in an assignment of the peak identity by means of computed membership values. A detailed explanation of the fuzzy algorithm and the logic of the procedure is given in refs. 8 and 9.

Fully automated method development can only be achieved with modules providing peak tracking and optimization embedded in an automated HPLC system consisting at least of an autosampler, pumps, a single- or multi-channel detector and a computer-based data station. For this purpose, commercially available HPLC equipment, as described in the Experimental section, was used with an MS-DOS-based data system, which includes the possibility of running a user program from system control level. It is possible to pass control from the data acquisition/reduction level to an interface file which allows communication with the above-mentioned software modules. This provides a means of transferring the integration data, containing a list of retention times, and area data from two-channel data acquisition/reduction, to the peak-tracking routine which assigns an identity to all detected signals. A retention data/peak assignment table is produced and used by the OPTIM program to construct the optimization hypersurface. The proposed best experimental conditions for the next experiment, as well as the least-density point of experimental conditions, are then transferred from the optimization module to the system level of the data station, and again chromatographically processed. Consequently, a loop is installed into a commercial HPLC system enabling automated method optimization within the limits of HPLC equipment that was originally not designed for this specific purpose.

EXPERIMENTAL

Instrumentation and software

The fully automated HPLC system was a Kontron (Zurich, Switzerland) liquid chromatograph consisting of four D-420 pumps, a D-460 autosampler furnished with a 100-μl Hamilton dosing syringe, two M-800 mixing chambers and a Model 400 column oven. The system was controlled by a Kontron 450 PC/AT data station. The detector was a Perkin-Elmer (Norwalk, CT, USA) LC-235 diode array detector with a flow-cell path length of 10 mm and a total volume of 4 μl, coupled to a Perkin-Elmer

GR-100 graphic printer–plotter. Throughout all experiments the signals were monitored at two wavelengths (210 and 230 nm).

The software for system control and data acquisition was commercial MT2 software (Softron Munich, Germany). It was interfaced with OPTIM software (having the optimization and peak-tracking routines) developed at the Graz University of Technology (Graz, Austria). For the three-dimensional representation of the optimized hypersurface, the Statgrafics program (STCS, Rockville, MD, USA) was used.

Reagents and solvents

Acetonitrile and *tert.*-butyl methyl ether (TBME) were of HPLC grade and obtained from Rathburn (Walkerburn, UK) and Fluka Chemie (Buchs, Switzerland). Water was deionized and filtered. Separations were carried out on an Ultrasphere ODS column, 25 × 0.46 cm I.D., 5 μm particle size (Beckman Instruments, San Ramon, CA, USA).

A synthetic mixture of eight different cyclosporins (A, B, C, G, L, T, Isocyclosporin A and dihydrocyclosporin A) was prepared in house. Their full structures have been published elsewhere [10].

DISCUSSION

The system described was designed for fully automated HPLC optimization. Four pumps allow ample possibilities for the use of different combinations of organic solvents and water. Furthermore, pH and ionic strength can be very easily changed by simple variation of the mixing ratio of different buffer solutions. A mixture of different cyclosporins (cyclic undecapeptides) was used as an example for the evaluation of the computer-aided optimization. Our goal was to optimize the isocratic HPLC method (mobile phase selectivity) and to evaluate its robustness.

Generally, cyclosporins have high solubility in diethyl ether, TBME, methanol and acetonitrile, while their solubility in water is very poor. The most common mode of separation of different cyclosporins is reversed-phase HPLC [11], with a mobile phase consisting of one of the binary mixtures methanol–water or acetonitrile–water. To achieve better selectivity we have applied a ternary mobile phase consisting of TBME, acetonitrile and water. On the basis of previous knowledge regarding miscibility of this three-component liquid phase (limited miscibility of TBME with water), the upper and lower boundaries of the solvent composition were selected appropriately.

In the instrumental set-up, the sum of the flow-rates of the pumps A and B was set to be 100%. The same rule was applied to pumps C and D. To start the optimization it was necessary to define the total flow-rate resulting from all four pumps and to input the percentage of the total flow for pumps B and D. The values for the pumps A and C were adjusted automatically. Pumps A and C were delivering acetonitrile, while B and D were supplying water and TBME, respectively. For the start of the optimization four runs were selected representing the boundary conditions of the preselected variable space. The restraints were set between 60 and 80% for pump B and between 10 and 30% for pump D.

It is possible to select among several predefined stop criteria for the optimization process. Once this measure has been accomplished the program stops. A resolu-

tion of defined value can be used as one possible criterion, another being defined as the maximum number of experiments to be performed. The latter can be used in order to stop the unit before the reservoirs of the mobile phase are emptied in automated overnight operation, *e.g.* Finally, the program will stop if convergence of the optimum conditions is predicted by the optimization program. For these experiments the criterion was set to stop the procedure when five consecutive optima had been proposed within the limits of ± 3% of the variable composition.

After the four initial isocratic runs (conditions supplied by the user) had been executed, the peak-tracking routine was automatically activated and performed. The identities of the components were traced as described above, based on area measurements at the two different wavelengths and the peak elution order. This has been found to provide a reliable method for peak tracking even at conditions of signal overlap. As described in the Introduction, the routine searches first for the pure peaks. Peaks not recognized as pure are then compared with the reference run. The run with the highest number of signals is automatically selected as the reference. Linear combinations of the non-assigned reference peaks are computed in order to detect the overlapping peaks and to identify properly the peaks in the experimental run. In Table I the results of the signal identification routine are reported. Upon its completion the program proceeds to the calculation of the optimization hypersurface.

The aim of this optimization software is to find the best chromatographic attribute, which is represented by the maximum value of the CRF_2. Furthermore, the optimization routine forces the experiments to step also into the region of the least available information, to characterize systematically the entire parameter space.

After the first four runs, the conditions which provide the highest value for the CRF_2 (calculated values for pumps B and D) were computed by OPTIM, and the result was automatically transferred to the control level of the data station in order to perform the next experiment. In addition to looking for the best conditions in the predefined space, this software also determined the point with the lowest density of information, *i.e.* the least density point of experiments. Consequently, after verification of the optimum, these conditions were executed. In this way, performing consec-

TABLE I

SIGNAL ASSIGNMENT AFTER EIGHT EXPERIMENTS WITH RUN NO. 3 AS REFERENCE RUN

mf is a membership function (see text).

Signal No.	Reference signal No.	Peak area		mf Value
		210 nm	230 nm	
1	1	458.89	243.86	0.843
2	2	220.20	123.22	0.844
3	3	660.77	361.63	0.844
4	4	312.76	170.06	0.844
5	5	554.82	299.24	0.844
6	6	80.41	42.59	0.845
7	7	82.20	46.72	0.833
8	8	494.00	280.08	0.845

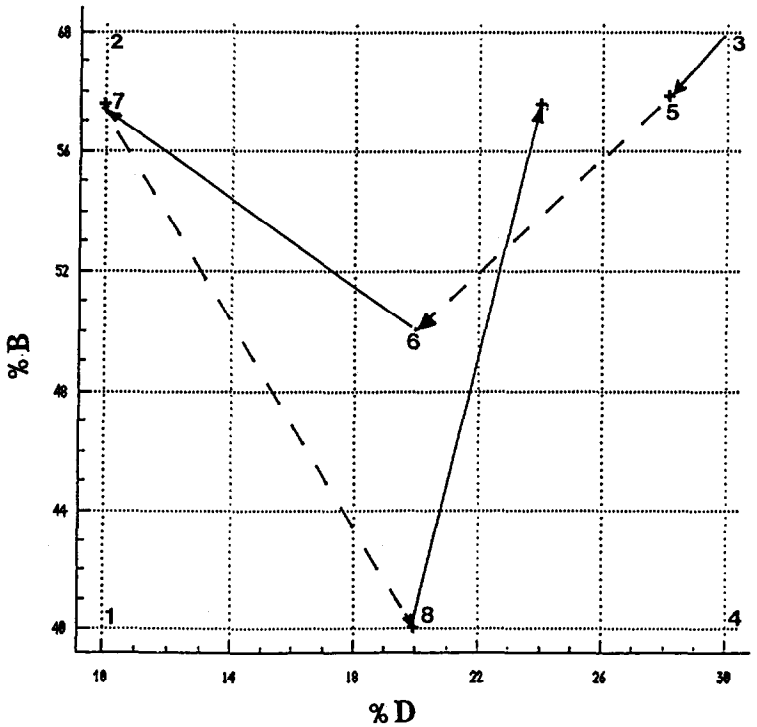

Fig. 1. Execution order of the eight optimized runs in the space defined with the percentage values for pumps B and D satisfying the conditions for a total flow for pumps A and B (A + B = 100%) and pumps C and D (C + D = 100%).

utively the best run and the experiment according to the least density of information, the program searches for a global and not for a local maximum.

Fig. 1 displays experimentally investigated points in order to visualize the order in which the experiments proceeded. As depicted in Fig. 1, after four experimental runs (labeled 1–4) experiment No. 5 was performed. This was carried out under the optimal conditions deduced from the initial four experiments. The new experimental point was used for the further optimization, but owing to the global scheme of this procedure the next chromatogram was performed under conditions corresponding to the least density point of experiments. The region of the least density information point, labeled 6, corresponds to the conditions B = 50%, D = 20%. The optimization routine then calculated new optimum conditions, marked as point 7, and the resulting experiment was executed. In a manner identical to that described above, the point with the least density information, No. 8, was performed and the refined optimum marked as + had a corresponding value of $CRF_2 = 0.9321$. To visualize more easily the development of the hypersurface, the numerical value of the CRF_2 which increases as the number of the performed experiments increases was shown in Fig. 2a, b, c and d with the hypersurfaces obtained after 4, 5, 6 and 8 performed experiments, respectively. The hypersurfaces are derived from the presently available experimentally obtained data points. Intermediate values are derived from the non-linear in-

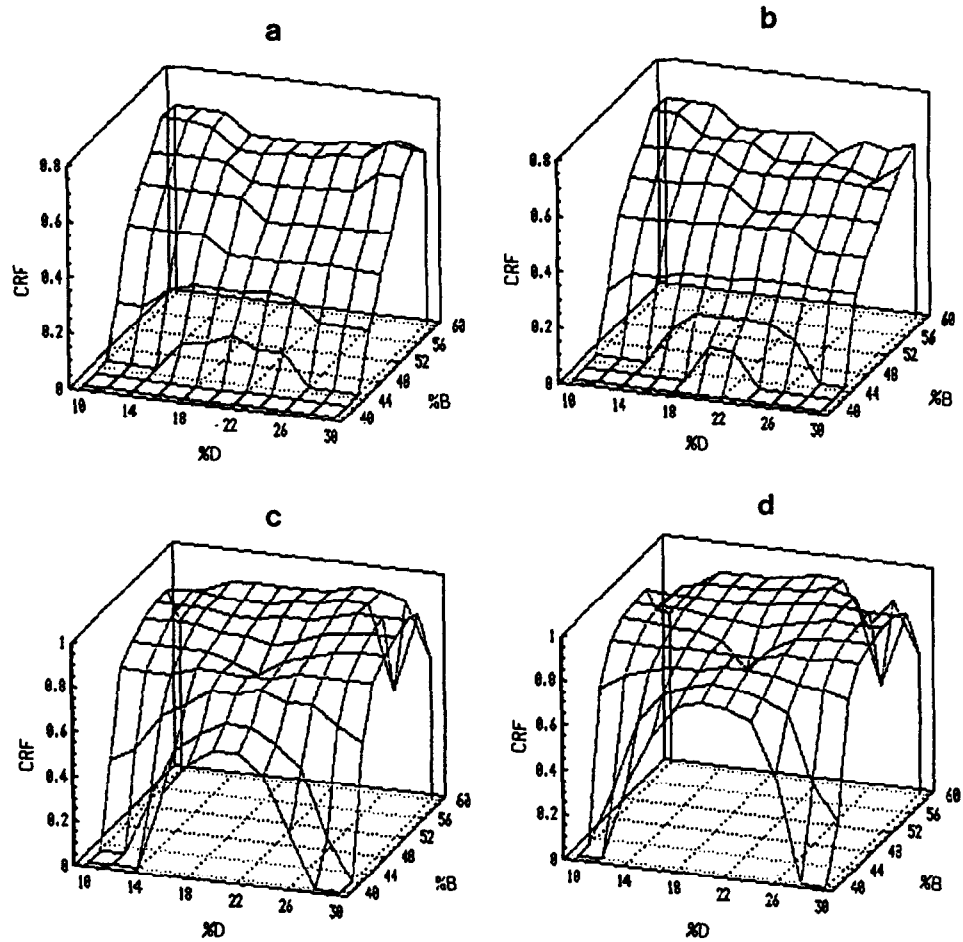

Fig. 2. Evolution of the hypersurface with increasing number of experimental points. (a) Surface derived from four experimental points; (b) five experiments; (c) six experiments; (d) eight experiments. *y*-Axis represents values for calculated CRF, while %D and %B are as in Fig. 1.

terpolation by means of the moving least-squares algorithm. It should be noted that for Fig. 2c and d, the CRF_2 axis is in the range 0–1, while for Fig. 2a and b it is from 0 to 0.8. The surface propagated to the region bounded with B = 48–60% after the five experiments. In Fig. 2c and d this increase in the CRF_2 is even more pronounced, and it is clear that the regions of previously low CRF_2 are becoming more important with increasing experimental knowledge (B = 40% and D in the region from 14–28%).

Once a proper separation quality has been accomplished, the next step is to evaluate the quantitative utilization of the procedure, *i.e.* the robustness of the chromatographic method. The robustness, being defined as a high plateau on the hypersurface, can be evaluated from the three-dimensional representation of the hypersurface. It should be emphasized that the first goal of this procedure was the optimization of the mobile phase composition with subsequent evaluation of the

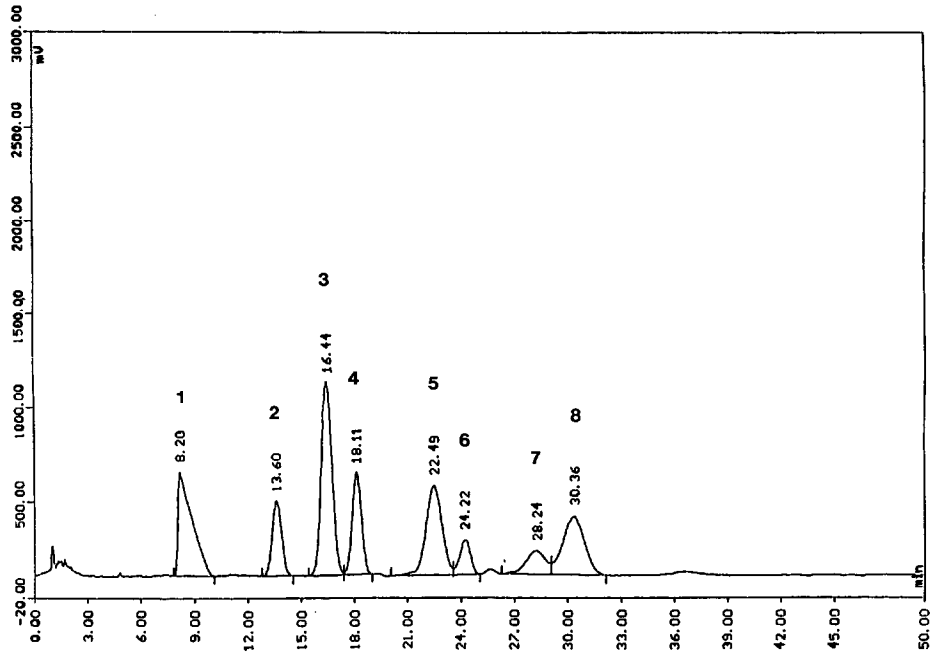

Fig. 3. Chromatogram acquired under the computed optimum mobile phase conditions: acetonitrile–TMBH–water (50.9:6.0:43.1). Total flow-rate was 2.0 ml/min. Column temperature: 353 K. Peaks: 1 = isocyclosporin A; 2 = cyclosporin C; 3 = cyclosporin B; 4 = cyclosporin L; 5 = cyclosporin A; 6 = dihydrocyclosporin A; 7 = cyclosporin T; 8 = cyclosporin G.

robustness of the newly developed chromatographic method. From Fig. 2d, it is possible to appraise the robustness of the developed method. The highest values for the CRF_2 and the best conditions concerning method robustness were limited to the region defined with the values for pump B from 52 to 60% and pump D from 16 to 26%. The experimental conditions with the highest CRF_2 were confirmed by a verification experiment under the proposed conditions (B = 57.5% and D = 24%) (Fig. 3). This demonstrates the interconnection between the quality of the chromatographic separation and the numerical value of the CRF_2. A chromatogram obtained under conditions yielding a $CRF_2 = 0.6$ is displayed in Fig. 4. It can be seen that the peak resolution is worse than that obtained under optimal conditions represented in Fig. 3. For $CRF_2 = 0$, the values for the pumps B and D were 40 and 30%, respectively. The chromatogram with overlapping peaks is presented in Fig. 5.

When the same experimental data (eight runs) were used as an input for the off-line version of the OPTIM program which includes time normalization, a new optimum was calculated at a mobile phase composition corresponding to B = 60% and D = 20%. The chromatogram in Fig. 6 clearly demonstrates that the selection of the quality criterion determines the definition of the optimum conditions. The inclusion of time normalization in the calculation of the CRF results in shorter analysis time, which may help to increase the productivity.

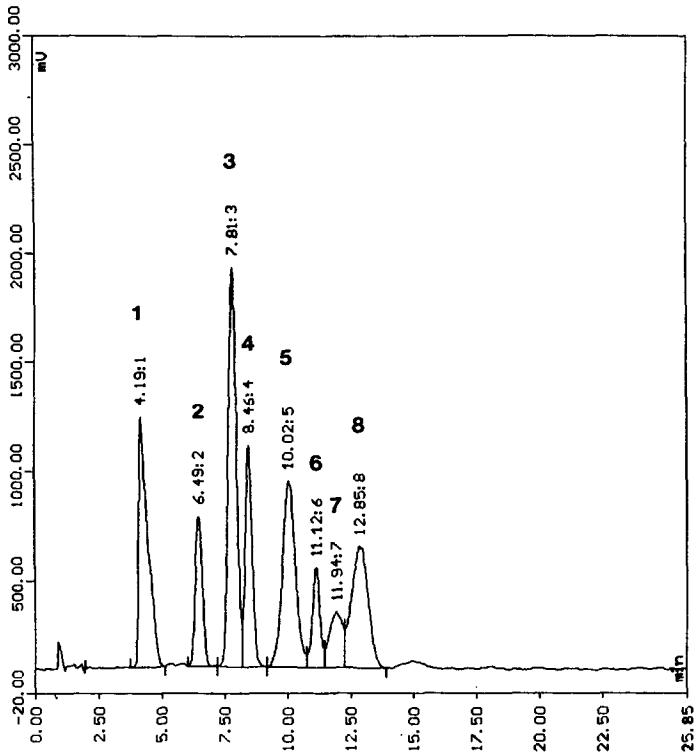

Fig. 4. Chromatogram obtained with the mobile phase conditions acetonitrile–TMBH–water (47.5:7.5:45.0). Other conditions as in Fig. 3. These conditions correspond to the point 3 in Fig. 1.

CONCLUSION

The optimization routine aimed to yield, within a precisely defined experimental framework, a preselected level of resolution for all pairs of components in a complex mixture. After an initial number of predefined experiments had been executed, the best mobile phase composition was ascertained using a strategy based on non-linear interpolation between measured retention data of the individual sample constituents and a subsequent computer construction of the optimization hypersurface. A software interface permits automatic transfer of the optimized values for the mobile phase composition to the pumps. Furthermore, to avoid reaching a local maximum, chromatographic experiments with solvent compositions representing the least density in space of variables were also performed systematically. An automated on-line identification of the solutes was accomplished using two-channel monitoring (UV absorbance at two different wavelengths) and using the recently introduced fuzzy peak-tracking approach. This peak-recognition routine was always active during the optimization process.

Robustness of the method can be judged from the optimization hypersurface. Owing to its empirical nature this approach can be applied to any chromatographic

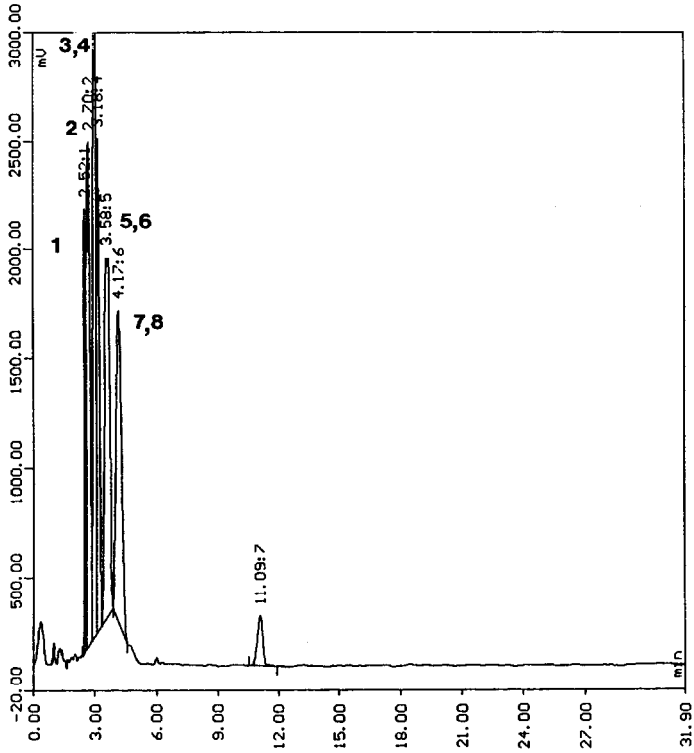

Fig. 5. Chromatogram obtained with the mobile phase conditions acetonitrile–TMBH–water (62.5:7.5:30.0). Other conditions as in Fig. 3. These conditions correspond to the point 4 in Fig. 1.

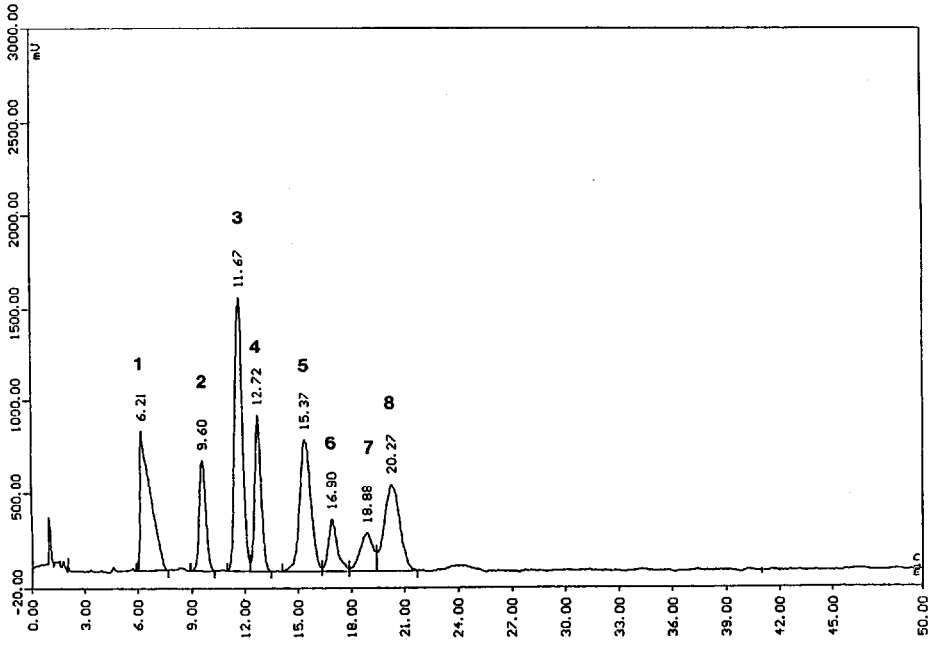

Fig. 6. Chromatogram obtained with the mobile phase conditions acetonitrile–TMBH–water (50.0:5.0:45.0). Other conditions as in Fig. 3.

separation process, even in cases where synergistic effects from separation variables are to be expected. The model is based on the real outcome of the experiment, and therefore makes it possible to obtain a more accurate interpretation on the robustness of the method. It is very flexible concerning the pick of the variables, since it does not rely on a mechanistic relationship between the chromatographic retention behavior of the sample constituents and variables, *e.g.* assuming linear or logarithmic dependence. This combination of the optimization routine, chromatographic data system and HPLC instrumentation offers a guide to higher productivity for analytical laboratories committed to HPLC method development and validation.

REFERENCES

1 P. J. Schoenmakers, *Optimization of Chromatographic Selectivity, (Journal of Chromatography Library*, Vol. 35), Elsevier, Amsterdam, 1986.
2 J. C. Berridge, *Techniques for the Automated Optimization of HPLC Separations*, Wiley, Chichester, 1985.
3 L. R. Snyder, J. L. Glajch and J. J. Kirkland, *Practical HPLC Method Development*, Wiley, New York, 1988.
4 E. P. Lankmayr, W. Wegscheider and K. W. Budna, *J. Liq. Chromatogr.*, 12 (1989) 35–58.
5 R. E. Kaiser, *Gas Chromatographie*, Geest & Portig, Leipzig, 1960, p. 33.
6 W. Wegscheider, E. P. Lankmayr and K. W. Budna, *Chromatographia*, 15 (1982) 498–504.
7 A. C. J. A. Drouen, H. A. H. Billiet and L. De Galan, *Anal. Chem.*, 57 (1985) 962–968.
8 M. Otto, W. Wegscheider and E. P. Lankmayr, *Anal. Chem.*, 60 (1988) 517–521.
9 E. P. Lankmayr, W. Wegscheider, J. Daniel-Ivad, I. Kolossvary, G. Csonka and M. Otto, *J. Chromatogr.*, 485 (1989) 557–567.
10 A. von Wartburg and R. Traber, in J. F. Borel (Editor), *Progress in Allergy*, Vol. 38, Karger, Basle, 1986, pp. 28–45.
11 M. Furlanut, M. Plebani and A. Burlina, *J. Liq. Chromatogr.*, 12 (1989) 1759–1789.