Journal of Chromatography, 485 (1989) 477–500
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 21 982

OBJECTIVE FUNCTIONS IN EXPERIMENTAL AND SIMULATED CHROMATOGRAPHIC OPTIMIZATION

COMPARATIVE STUDY AND ALTERNATIVE PROPOSALS

R. CELA*
Analytical Chemistry Department, Faculty of Sciences, University of Santiago de Compostela, Santiago de Compostela (Spain)
and
C. G. BARROSO and J. A. PÉREZ-BUSTAMANTE
Analytical Chemistry Department, Faculty of Sciences, University of Cádiz, Cádiz (Spain)

SUMMARY

The rôle of the objective function in chromatographic optimization and method development is emphasized by a critical study of several functions proposed in the literature. Different properties are discussed for characterizing an ideal (if this exists) chromatographic objective function to be used in experimental or simulated off-line chromatographic optimization. In this context a new information theory-based criterion is proposed.

INTRODUCTION

As liquid chromatography has become acknowledged as an increasingly powerful separation technique, it has been faced with the resolution of more and more complex and challenging problems. The conventional approach, which relies on the expertise and intuition of the chromatographer as the only means of obtaining good separations, has revealed its shortcomings when faced with this kind of problem. Formal optimization strategies are currently used, providing more objective optimization of complex separations and the possibility of computer control.

In the more general case, the goal of the optimization process must be to improve the separation between all the peaks representing the individual components of a particular mixture (sample), obtaining a chromatogram in which each peak will correspond to one (and only one) component with the condition that no significant overlapping between peaks takes place. On the other hand, it must be realized that an improvement in separation implies an increase in resolution between peaks, but the peak capacity is reduced if the resolution is increased without increasing the number of theoretical plates. Hence, although quantitative information is gained, qualitative information is lost, and therefore it is necessary to reach a compromise between the two. In the conventional approach it is the chromatographer who, more or less

0021-9673/89/$03.50

empirically, finds this compromise. In contrast, in formal optimization strategies a numerical criterion is needed to guide and stop the optimization process. This paper focuses on the role of the objective function in the development and optimization of chromatographic procedures.

FEATURES OF AN IDEAL CHROMATOGRAPHIC OBJECTIVE FUNCTION

An ideal chromatographic objective function must fulfil the following two fundamental requirements: (1) to have an effective means of comparison and differentiation of chromatogram quality and (2) to have an effective means of quantitative scaling of chromatogram quality. Moreover, several other features are necessary, or at least convenient: (3) to serve effectively the aims of the chromatographer; (4) to be affected by the controllable parameters in the hands of the chromatographer and not by the uncontrollable parameters; (5) to exhibit an understandable correlation with the controllable separation parameters in order to indicate to the chromatographer in a straightforward manner the way in which to improve the result of the next trial; and (6) not to suffer from mathematical limitations or inconsistencies.

These additional features are, in practice, very difficult to obtain. The large number of objective function (*OF*) formulations that have appeared in the literature during the last 15 years is a demonstration of this difficulty. This large number of proposals has been the subject of extensive discussion in recent books[1–3].

Condition 3 is, in fact, nearly impossible to achieve, because the aims of the chromatographer vary according to the analytical problem encountered and the chromatographic technique chosen to solve it. The resolution of all the peaks composing the mixture is, in practice only a particular situation. On many occasions a reduced number of components in the mixture is of interest to the analyst. Obviously, the strategies used to optimize both types of separations cannot be the same and a unique *OF* can hardly meet the goals of both types of strategies. The question to be answered in each instance is what the chromatographer thinks is a good (or at least acceptable) separation. This definition varies widely from the situations where standards of pure solutes are available to those (*e.g.*, in gel permeation chromatography) where standards (polymers) are composed of a more or less narrow range of molecular sizes, so the peaks always represent solute mixtures.

Therefore, it seems unavoidable to have *OF*s lacking general analytical validity, but serving specific goals for a particular problem. This fact was mainly responsible for the tendency to develop *OF*s on an *ad hoc* basis, as pointed out by Wegscheider *et al.*[4], and the introduction of user-selectable weighting factors in the mathematical formulation of most *OF*s, trying to gain performance in connection with any particular experimental variable. In fact, the use of weighting factors makes it easier for the chromatographer to adapt a particular *OF* to this particular problem. This approximation has proved to be very useful in many circumstances, but obviously introduces a significant subjectivity in the *OF*. Depending on the chromatographer's choice, very different results can be obtained for the same separation problem[5].

Let us consider as an example the *OF* proposed by Berridge[6]:

$$OF = \sum^{L} R_i + L^{w_1} - w_2 |\, T_A - T_L \,| - w_3(T_1 - T_0) \qquad (1)$$

where

R_i = resolution between the ith and the $(i + 1)$th peaks (if the measured or calculated resolution is >2.0, then $R_i = 2.0$ in the calculation);

L = number of peaks appearing in the chromatogram;

T_A = maximum acceptable time;

T_L = time to complete the elution;

T_1 = retention time for the first-detected peak;

T_0 = minimum desirable retention time for the first peak in the chromatogram;

w_1–w_3 = operator-selectable weighting factors which are usually[6] set in the range 0–3.

With w_2 and w_3, it is obvious that the time needed to complete each separation can play a very important role in routine analysis and, depending on the specific mixture to be separated (and the number of components of real interest), a poorer separation (from the point of view of the resolution) is sometimes preferred. Under these circumstances, the w_2 weighting factor must be set at the top of the range. On the other hand, w_3 mainly controls the resolution between the solvent front and the first peak of real interest. Depending on whether this first peak of interest coincides or not with the first-eluting peak in the mixture, the w_3 weighting factor must take different values in the prescribed range.

However, the real problem with this OF relies on the w_1 weighting factor. Overlooking the time weightings (*i.e.*, $w_2 = w_3 = 0$), Berridges's OF is reduced to the sum of two terms:

$$OF = \sum_{1}^{L} R_i + L^{w_1} \tag{2}$$

Since for the practical application of this criterion the resolution between adjacent peak pairs is limited to a maximum value of 2.0, for well resolved mixtures equation 2 gives

$$OF = 2(L - 1) + L^{w_1} \tag{3}$$

From this equation, it is obvious that the weight of the second term in the total value of the OF increases with increasing w_1 (as expected), but also with increasing peak number when $w_1 > 1$. When $w_1 = 1$ the weight of this second term is approximately constant and for values of w_1 approaching zero this term loses importance with increasing peak number.

The consequence is that the resolution term lacks relevance even for simple (small number of peaks) separations when $w_1 > 2$. For chromatograms of medium complexity the OF means simply counting the appearing peaks. On the other hand, if $w_1 = 0$, only the resolution is relevant in the final OF value. Thus, in practice values of w_1 between 1 and 2 seem to be the normal choice (Berridge[7] recommends the use of a value of $w_1 = 2$).

Condition 4 becomes more or less critical as a function of the formal optimization strategy used. In fact, some of them suffer from noise much more than others.

Condition 5 determines to a great extent the mathematical formulation of the

OF. In fact, as the number of criteria involved in the *OF* increases, it becomes more difficult to interpret the meaning of the final numerical value obtained in terms of the controllable parameters. For this reason the number of criteria included in most *OF*s has been limited to two (peak resolution or peak separation and time).

In some instances, the *OF* formulation represents a very particular situation, so the way in which this *OF* guides the optimization process is not necessarily the one which the chromatographer wants, but the condition forced by the mathematical formulation itself. In this instance the correlation with a particular set of values of some controlled parameters could be overemphasized.

Let us consider, for example, the resolution product criteria proposed by Schoenmakers and co-workers[1,8,9]. In chapter 4 of his book[1], Schoenmakers gives an example of a very simple separation (see Fig. 1, showing the corresponding simulated chromatogram for three different situations). The calculations using the normalized resolution product and the callibrated normalized resolution product,

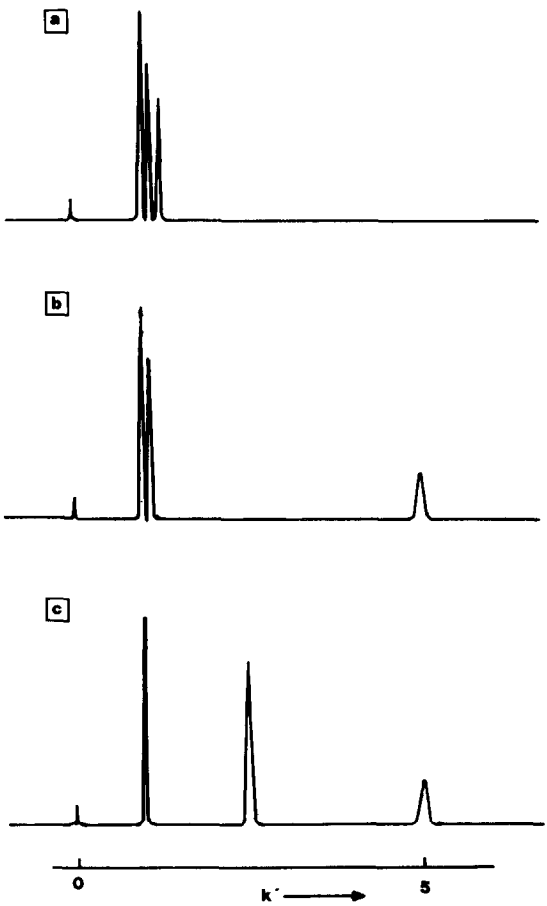

Fig. 1. Three simulated chromatograms. Constructed for $N = 10\,000$. Capacity factors are listed in Table II. (Taken from ref. 1).

$$NRP = \prod_{1}^{n-1} (R_{i,i+1}/\bar{R}) \tag{4}$$

$$CNRP = \prod_{0}^{n-1} (R_{i,i+1}/\bar{R}^*) \tag{5}$$

where

n = number of peaks;

$\bar{R} = [1/(n-1)] \sum_{i}^{n-1} R_{i,i+1}$

$\bar{R}^* = (1/n) \sum_{0}^{n-1} R_{i,i+1}$

lead to the conclusion that chromatogram c is the best one. Let us consider this case with some variations of the peak position (assuming a constant plate number) in chromatogram c. Table I summarizes all the data and calculations, including those given by Schoenmakers on page 137 of his book[1].

Case c2 is obtained by approximation of the first peak to the solvent front, but keeping constant (with regard to case c) the resolution between the three peaks included in the mixture to be separated. The result obtained for the *CNRP* criterion, however, is far worse than in case c. Nevertheless, the resulting chromatogram should be better, because it has been obtained in a shorter time ($k'_w = 2.3$ compared with $k'_w = 5$ in case c). This is one of the reasons argued by Schoenmakers for introducing time factors in these *OF*s.

However, time factors do not solve completely the drawbacks of these functions, as can be seen from Table I. In case c2 it is not possible to argue that the separation of the first peak and the solvent front (peak zero in this criterion) is deficient ($R_{0,1} = 2.38$); case c4 where $R_{0,1} = 10$ (and where the time factors have no influence) gives similar conclusions in this sense.

Case c3 assumes approximation to the solvent front only for peak 1 (hence time factors are irrelevant with regard to case c), and the resolution between peaks of interest improves. However, the criterion value in this case is again very poor. Case c4

TABLE I

SOME CALCULATIONS USING THE *CNRP* CRITERION

See Fig. 1 for chromatograms a, b and c. The calculations for these cases were taken from ref. 1, p. 137.

Case	k_1	k_2	k_3	$R_{0,1}$	$R_{1,2}$	$R_{2,3}$	$\bar{R}$	*CNRP*
a	1	1.1	1.25	16.6	1.22	1.72	6.51	0.13
b	1	1.1	5	16.6	1.22	24.1	13.97	0.18
c	1	2.5	5	16.6	13.6	13.2	14.5	0.98
c2	0.1	0.92	2.3	2.38	13.6	13.2	9.7	0.46
c3	0.1	2.5	5	2.38	26.1	13.2	13.9	0.30
c4	0.5	2.5	5	10.0	20.0	13.2	14.4	0.88
c5	1.2	2.5	5	18.7	11.4	13.2	14.4	0.93
c6	2	5	10	25	16.6	14.7	18.8	0.92
a2	0.05	0.1	0.18	1.22	1.16	1.75	1.37	0.95

is very similar to case c3 but now the approximation of the first peak to the solvent front is smaller. In this instance, the *CNRP* result is much better (compared with case c3), although not as good as case c.

Case c5 assumes that the first peak goes slightly further from the solvent front. The resolution is still excellent, although $R_{1,2}$ was slightly smaller than in case c. The *CNRP* value obtained was again smaller than case c.

Finally, case c6 assumes duplication of the capacity factors of all the peaks. In this instance, the chromatogram will take much more time to be developed. The *CNRP* value is slightly smaller than in case c, but curiously very similar to that obtained in case c5.

Therefore, it seems that chromatogram c satisfies the *CNRP* mathematical requirements very well (similar resolution between peaks and the solvent front). In fact, the poor result obtained for case a is mainly due to the separation between the first peak and the solvent front ($R_{0,1} = 16.6$) which gives a high value of $\overline{R}^*$. If the entire chromatogram approximates the solvent front (case a2), a very high value of the *CNRP* criterion can be obtained. Obviously, requirement 2 of the ideal *OF* is not fulfilled by this criterion.

Finally, failures in condition 6 could be the result of attempts to fulfil the other conditions previously mentioned. This is the case, for example, with the *CRF* proposed by Watson and Carr[10] or the *COF* proposed by Glajch *et al.*[11]:

$$CRF = \sum_{1}^{n} \ln(P_i/P_0) + w(T_A - T_L) \tag{6}$$

$$COF = \sum_{1}^{n} A_i \ln(R_i/R_{id}) + w(T_A - T_L) \tag{7}$$

where

P_0 = desired peak separation;
P_i = peak separation for the ith pair of peaks;
R_{id} = desired resolution for the ith pair of peaks;
A_i = weighting factor for the ith pair of peaks.

The other symbols have the same meanings as in previous equations.

Some of the mathematical drawbacks in these *OF*s are easily appreciated and have been pointed out in the literature[1,12–14]. In logarithmic-type *OF*s when there is total overlap between two or more peaks, $P_i = 0$ (or $R_i = 0$) and, therefore, the function tends to minus infinity. In this type of function that approaches zero at the maximum, no indication is given of whether a large number of peaks have been well separated or whether just a few peaks are poorly separated. In fact, when overlapping of a pair of peaks increases, the "disappearance" of a peak in the chromatogram apparently improves the value of the *OF*, provided that this is reduced. Because of this, compensatory terms must be introduced:

$$CRF = \sum_{1}^{m} \ln(P_i/P_0) + \sum_{1}^{n-m} \ln(0.01/P_0) + w(T_A - T_L) \tag{8}$$

(ref. 13) or

$$CRF = \sum_{1}^{n-1} \ln P_i - 100(n - m) \tag{9}$$

(ref. 15). In any case, these compensatory terms, although improving considerably the *CRF* behaviour, have an obvious empirical character, so the problems concerning the choice of the most adequate values or the general validity of the values proposed in the literature still remains.

The conclusion is that at present no single *OF* meets all the requirements needed to consider it as an ideal *OF* and, therefore, the field remains open for new and improved proposals.

OBJECTIVE FUNCTIONS BASED ON INFORMATION THEORY

The use of information theory with the aim of optimizing chromatographic separations dates from the papers by Massart and co-workers[16,17]. They questioned the usefulness of the resolution-based criteria as these only allow the quality of separation between pairs of peaks to be calculated, and do not give a general view of the quality of a multi-component separation. In theory, the informational objective functions can give this general view and so provide a better understanding of the purpose of chromatography, *i.e.*, to obtain information.

The optimization of the separation is achieved by maximizing the informing power, defined as

$$P_{\text{inf}} = \sum_{1}^{n} (\log_2 S_i) \tag{10}$$

where S_i is the reciprocal precision defined by Kaiser[18].

In the study by Massart and Smits[16], a series of equations were developed which were recognized as being of no practical application, as most of the parameters considered were known only for a few systems. Nevertheless, eqn. 10 was used in a paper published in 1975 by Smits *et al.*[17], centred on optimization through the simplex algorithm of the separation of five metallic ions by ion exchange, the precision being formulated as

$$S_i = 1/(\beta_{i-1,i} + \beta_{i,i+1}) \tag{11}$$

where β is the fractional overlap between peaks. The example used in this instance was too simple to judge the general validity of the approximation. In any case, problems due to peak cross-over were detected, and especially in the way the authors introduced the factor time in the calculation of P_{inf}. As this appears as a ratio, it implies that faster separations lead to the same value of P_{inf} than better, although slower, separations. On the other hand, the practical measurement of overlap presented difficulties not only at the time when the study by Smits *et al.* was published, but also at present, especially when a good mathematical description of the peaks is not available, as has been pointed out by Wegscheider *et al.*[4].

In 1980, Spencer and Rogers[19] used the same ideas, although with a slightly different calculus approach, to propose an objective function which they called separation number (*SN*) and applied it to gas chromatography. Basically, the *SN* algorithm uses the same equations and concepts as information theory applied to each datum from a digitized chromatogram. The amount of information derived from an experimental point on a chromatogram depends on the approximation with which this point coincides with its value predicted on the basis of data obtained for previous points. The amount of information provided by this point can take any value between 0 and 1. The *SN* represents the sum of the values calculated for each point in the chromatogram. In comparison with other objective functions, the advantage of *SN* is that it is not necessary to know the number of peaks, but it does not solve the problems derived from peak cross-overs and, in particular, has the disadvantage that the answer obtained depends markedly on scanning of the chromatogram, as stated by Debets *et al.*[14], and lacks general applicability as acknowledged by the authors. Recently, Van Hare and Rogers[12] published a new objective function which moves away from the informational *SN* criteria.

Since then, no new proposals have appeared in the literature, although Eckschlager and Stepanek[20] published a paper towards the end of 1982 concerning the applications of information theory in analytical chemistry, supporting its applicability to the optimization of high-performance liquid chromatographic separation processes, although no specific references were given.

Schoenmakers[1] stated that at present the fractional peak overlap seems to be merely a theoretical proposition, owing to the difficulties encountered in the accurate and realistic calculation of this criterion. We have developed an information theory-based criterion (*IAC*) by extrapolating the theoretical proposals of Liteanu and Rica[21] for thin-layer chromatography, which can be used in simulated off-line chromatographic optimization. This criterion has been included as an option in the PREOPT package[22].

The problem of the optimization of a chromatographic separation of any kind can be posed assuming that before carrying out the separation we have a group of n substances, their existence in the sample being equally probable:

$$X = (x_1, x_2, \ldots, x_i, \ldots, x_n) \tag{12}$$

A chromatographic experiment can then be carried out to determine which of the n possible substances really exist in the sample on the basis of retention time measurements of each peak. This is obviously a typical proposal in qualitative analysis which can be adapted to our problem because when two or more peaks appear strongly overlapped in the chromatogram (which happens when the separation has not been optimized), this is equivalent to the absence (impossibility of detection) of one or several peaks. Therefore, the quality of a chromatogram is directly related to the number of species that "do not appear" in it (owing to overlapping) and, finally, to the qualitative aspect of the experiment. In theoretical models it is normal to accept that retention time measurements which can be used to identify the species are affected by random errors according to a Gaussian distribution. It can also be accepted that the standard deviation of all the peaks may be analogous or not, which is the same as considering all the peaks as being of similar width (*e.g.*, gradient separations), or variable as a function of time (isocratic separations).

Assuming a certain quantification interval (q_{tr}), *i.e.*, the interval on the time scale according to which the chromatogram will be sampled, the distribution of probabilities before the experiment is performed will be:

$$P(X) = \begin{Bmatrix} x_i \\ p(x_i) \end{Bmatrix} = \begin{Bmatrix} x_i \\ 1/n \end{Bmatrix} \tag{13}$$

provided that all the components are equally probable and in fact are known to be present in the sample. The entropy [$H(X)$] of this distribution will be a measurement of uncertainty before the experiment, which can be calculated according to the equation

$$H(X) = -\sum_{1}^{n} p(x_i) \log_2[p(x_i)] = \log_2(1/n) \tag{14}$$

Obviously, it is necessary to know beforehand the number of species to be separated in order to calculate $H(X)$.

The set of events corresponding to the domain of the results of measurement for a quantification margin q_{tr} will be

$$Y = (y_1, y_2, \ldots, y_j, \ldots, y_m) \tag{15}$$

where $m = t/q_{tr}$, t being the time necessary to complete the chromatogram and corresponding to the intervals where the presence of a peak is detected or not. Within the limits we have established for the distribution of results (tr), the probabilities of conditional events (y_j/x_i) (*i.e.*, the peak corresponding to the species x_i appears in the interval y_j of the time scale) are simply evaluated by means of normal distribution values. Thus, for example, the graph in Fig. 2 shows the calculation and significance of the conditional probability $p(y_j/x_i)$, corresponding to the shaded area in the graph. It is therefore possible to evaluate a matrix of m rows and n columns that make up the conditional probabilities matrix $p(y_j/x_i)$ for the system.

The probabilities $p(y_j)$ for events in group Y after the experiment is finished are evaluated from the previous matrix according to the equation

$$p(y_j) = \sum_{1}^{n} p(x_i)\, p(y_j/x_i) \tag{16}$$

and, finally, the matrix of conditional probabilities $P(x_i/y_j)$ is calculated according to the equation:

$$p(x_i/y_j) = \frac{p(x_i)\, p(y_j/x_i)}{\sum_{n} p(x_i)\, p(y_j/x_i)} \tag{17}$$

a

$$p(y_j/x_i) = 1/\sigma(2\pi)^{1/2} \int_{y_j - (q_{tr}/2)}^{y_j + (q_{tr}/2)} \exp.\,[(y-x_i)^2 / 2\sigma^2]\, dy$$

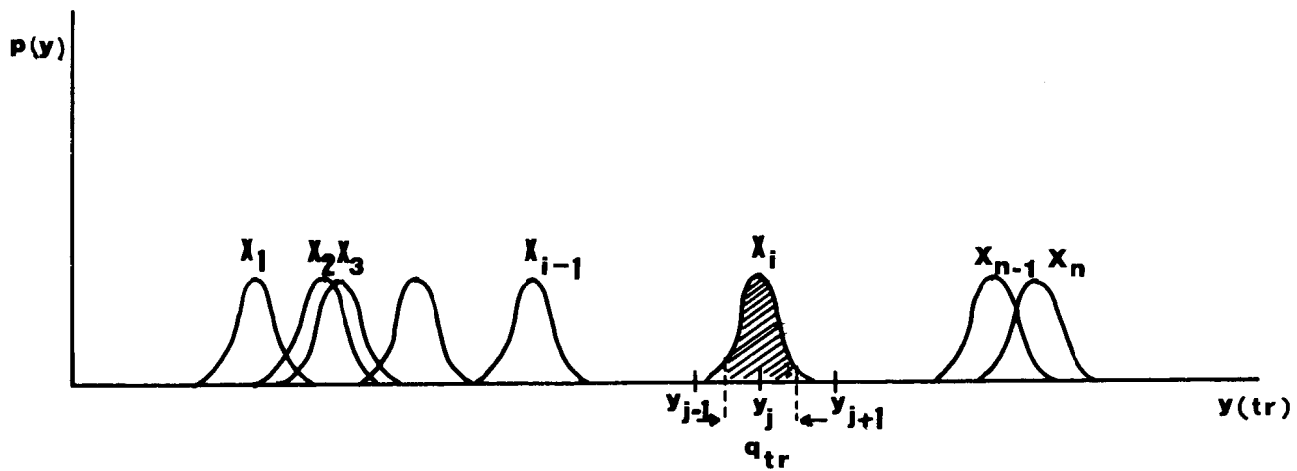

Fig. 2. Schematic view of *IAC* criterion calculation as applied to a chromatographic separation.

whereby a new matrix, m by n, is thus obtained, which in turn using the equation

$$H(X/y_j) = -\sum_{1}^{n} p(x_i/y_j) \log_2 p(x_i/y_j) \tag{18}$$

leads to the uncertainty (entropy) after the experiment is over (chromatographic separation), which subtracted from the initial entropy (eqn. 14) gives the amount of information $I(X/y_j)$ provided for each quantitation interval, while the average amount of information $[I(X/Y)]$ from the experiment is obtained from the following equations:

$$I(X/Y) = H(X) - H(X/Y) \tag{19}$$

$$H(X/Y) = \sum_{1}^{n} p(y_j) H(X/y_j) \tag{20}$$

In a schematic manner, Fig. 2 depicts this calculation process. For a given chromatogram (Fig. 2a) (assuming $n = 8$ peaks and a quantification interval of about half of the peak width), the evaluation of the areas corresponding to each peak along the axis divided into twenty quantification intervals gives a matrix $[A] = p(y_j/x_i)$ (Fig. 2b).

Each column of this matrix defines the probability of finding the peak x_i along the y (time) axis, so that in each column all the values are zero except for those intervals where the peak x_i is detected. Consequently, each row gives information about the number of peaks contributing to the chromatogram profile for each quantification interval. Obviously, the number of elements in each column common to two or more peaks gives a quantitative measurement of the relative overlap between those peaks. The aim of the information amount criterion (IAC) is to evaluate in terms of information units this overlap between peaks along the time axis. With this aim a new matrix $[B] = p(x_i/y_j)$ is calculated (Fig. 2c) summing all the elements in each row and dividing each of the elements of this row in matrix $[A]$ by the sum obtained (see eqn. 17). Obviously, if in a row (y_j) of the matrix $[A]$ the contribution of only one peak is present, the sum of this row will be equal to $p(y_j/x_i)$ and the corresponding element in matrix $[B]$ will be equal to 1 (*i.e.*, peak 6). In contrast, if more than one peak contributes to the chosen row in matrix $[A]$, the elements corresponding to this row in matrix $[B]$ will be lower than 1. If the matrix $[B]$ finally obtained contains one or more columns composed only of zeros and ones, this means that the peak represented by this column could be accurately quantified because no overlap with other peaks takes place. Hence the optimum separation will be represented by a matrix $[B]$ in which all the columns contain only zeros and ones. During the optimization process the goal is to evaluate the contribution of the columns having elements different from zero or one. This is achieved by calculating and adding up the values of eqn. 18 for all the intervals in which the time axis was split up. In this way, the magnitude of $H(X/Y)$ is calculated and from eqn. 19 the information amount is obtained.

The magnitude $I(X/Y)$ can be used as an objective function and represents the global quality of the chromatogram, provided that on resolving all the peaks each quantitation interval gives a maximum of information and the average amount of information thus obtained will be equal to the initial uncertainty, which will then be

eliminated. The criterion so defined can take any value between 0 and $H(X)$, which assumes the upper limit, provided that once all the peaks are resolved to the baseline a further separation between them will not give additional information. Obviously, a time limit must be considered in order to avoid unnecessary long chromatograms. The magnitude $H(X/Y)$ can be used similarly as an objective function within the $H(X) \rightarrow 0$ range because when all the peaks have been separated from each other the uncertainty after the experiment is non-existent.

DEVELOPMENT AND IMPLEMENTATION OF THE *IAC* CRITERION

As we have stated before, the *IAC* criterion was developed to be used in the PREOPT package. This simulation off-line optimization package has been described elsewhere[22]. However, since the practical way in which the *IAC* works has not yet been published, a few important implementation characteristics will be described here.

The *IAC* criterion is applied to evaluate simulated chromatograms in which the retention time of the peaks is obtained by means of the "step model" (which is the basis of the PREOPT) and their width is calculated according to the equations given by Jandera and Churáček[23] for isocratic or gradient elutions. The *AIC* in its developed form does not require peak identity assignment, as it is developed from a typical qualitative scheme. In fact, the criterion measures the number of peaks appearing in the chromatogram and assimilates them to a normalized probability distribution by measuring the existing overlap between the Gaussian curves. Therefore (assuming a constant flow-rate of the mobile phase and no changes in the column), the only previous data required are the number of peaks and the plate number of the column used. It is obvious that if two completely overlapping peaks appear it is impossible to detect which ones are overlapping and in which positions (unless other qualitative data are available). This does not pose any special problem because as far as the calculation of the conditional probabilities matrices is concerned, it does not matter if a given peak overlaps with one or another peak, the only question that matters is the extent of the overlap.

On the basis of this consideration, the peaks that for one reason or other are not detected in the chromatogram (either as a result of complete overlapping or as derived from being eluted outside the prefixed time limit) were arbitrarily assigned a retention time identical with that of the last-detected peak, so that the calculation is not affected. However, although the work is done with simulated chromatograms (in which by definition all the peaks exist and their positions can be easily found; the computer has this information), a much more realistic situation is obtained when the peaks that overlap strongly are removed from the chromatogram. For instance, in Fig. 1 peaks 2 and 3 show extensive overlap and will be hardly distinguishable. Therefore, in a realistic chromatogram of this mixture only seven peaks will appear.

Another consequence of this operation is that peak cross-over is ignored if the criterion is used in an optimization process. Obviously, a routine is needed to decide when two peaks are indeed indistinguishable or not. A typical second-derivative valley search routine or the use of graphical criteria according to Snyder and Kirkland's data[24] can be used with this aim.

In any event, the appearance of shoulders should be considered as a special case and, if necessary, their contribution to the *IAC* should be calculated. Generally, two

types of situations have been considered which have been referred to as the appearance of a total shoulder (in the case where a valley, however small, cannot be detected but the second peak is clearly detectable) and what has been called a partial shoulder (the case in which the valley, however small, can be detected).

The calculation routine of the *AIC* includes specific counters for the two situations in such a way that once the uncertainty is calculated after the experiment, the value obtained is increased as a function of the number of total or partial shoulders that appear in the chromatogram. However, in most instances the use of these penalizations does not improve the performance of the criterion, so it can be omitted.

On the other hand, one of the main problems of the *IAC* routine is the establishment of the quantification interval. This interval must be small enough to permit discrimination of narrow peaks not strongly overlapped. As can be seen in Fig. 3a, if a wider interval is chosen both peaks must be considered as totally overlapping by the criterion because in the interval y_j both are fully integrated. Hence in this situation the criterion will fail in the evaluation of the quality of the chromatogram. On the other hand (Fig. 3b), if a narrow interval is chosen the accurate discrimination between peaks will be achieved, but the number of intervals considered in the entire chromatogram increases dramatically and so does the time needed to make the necessary calculations and the computer memory needed to store the $[A]$ and $[B]$ matrices. For instance, considering a column having 3000 theoretical plates, peaks appearing at the beginning of the chromatogram will have widths of about 0.05–0.1 min. On the other hand, peaks appearing at the end of the chromatogram (*i.e.*, $k' = 10$) will have widths of about 0.8–1.0 min. If we define an interval of 0.05 min as ensuring that the first peaks are well detected by the routine, a usual complete chromatogram (*i.e.*, 20–30 min) will have a total of 500 intervals so we must manipulate two matrices of 500 rows by a number of columns equal to the number of peaks. With complex mixtures this needs a lot of memory and calculation time.

Moreover, with this approximation the earlier peaks will be integrated whole by in only one interval whereas the last peaks in the chromatogram will be shaped by 20–30 intervals. Obviously, the possibilities of discriminating between adjacent overlapping peaks are greater in the final region of the chromatogram. However, if we reduce the width of the interval (*i.e.*, 0.01 min) to give a good shape of the earlier peaks, the memory and time needed increase by a proportional factor, making it very difficult or impossible to manipulate the matrices with a microcomputer.

Consequently, it would be better not to use a fixed value for the quantification interval as we have done in Fig. 3, but an interval of increased width with increasing elution time. In this way, we can choose an initial interval that is sufficiently narrow accurately to integrate and discriminate earlier peaks, but all the peaks in the entire chromatogram will be shaped by an approximately equal number of intervals, so that the discriminatory power of the *IAC* criterion is constant despite the chromatogram region evaluated. In practice, it is desirable that each peak be shaped in matrix $[A]$ (in fact, matrix $[B]$ is used in a reduced form in the package) at least by three to four intervals so that peaks that do not overlap very much can be accurately evaluated. This implies the use of an interval width of about $\sigma(x_i)$ [where $\sigma(x_i)$ is the standard deviation of the gaussian peak x_i to be evaluated], so that the interval must be defined as a function of the plate number of the column,

$$q_{\mathrm{tr}_i} = q_{\mathrm{tr}_{i-1}} + 1/N^{\frac{1}{2}} \tag{21}$$

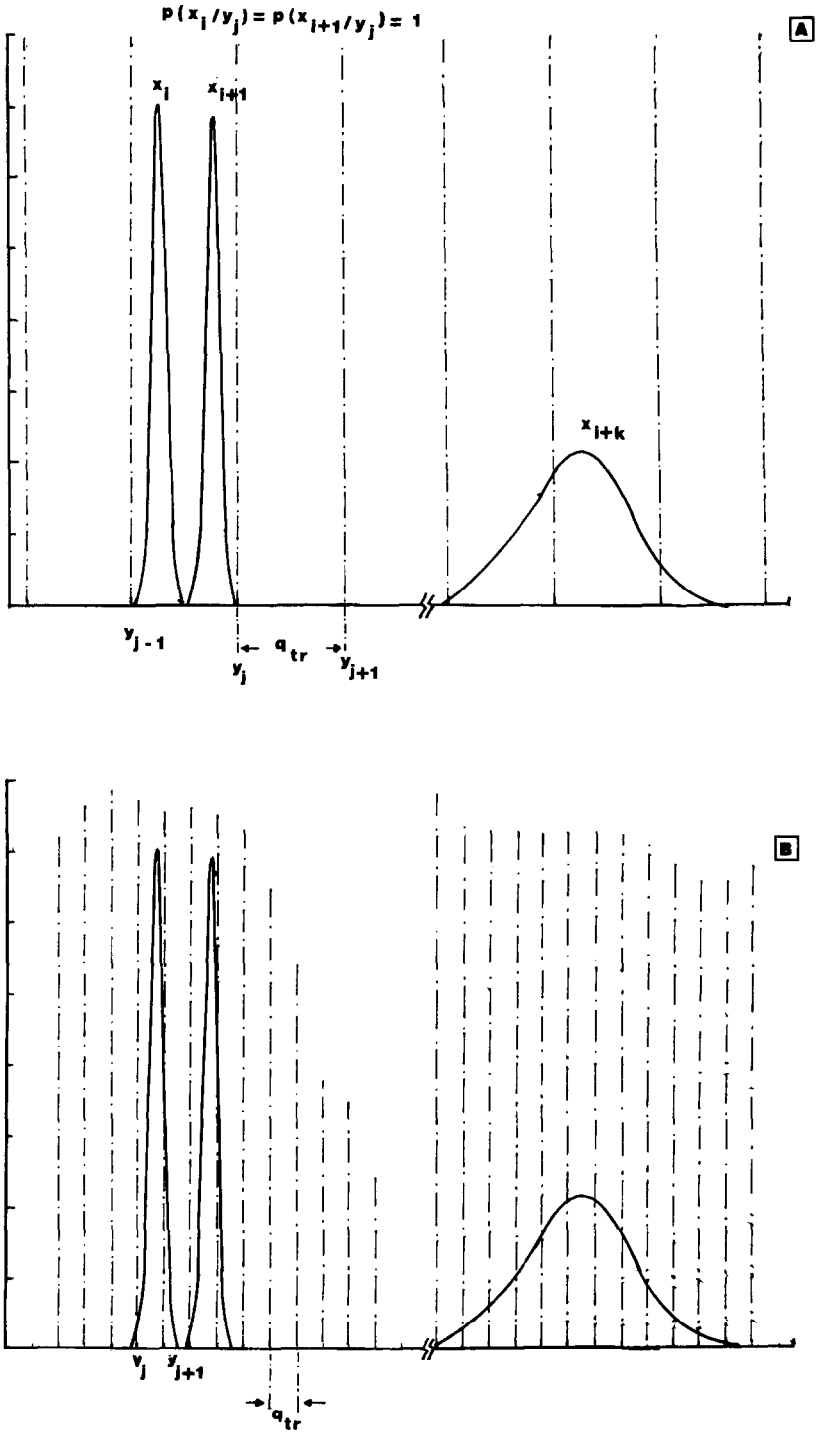

Fig. 3. Influence of the quantification interval (q_{tr}) in the performance of the *IAC* criterion.

In this way, the quantification interval thickness increases in the same way as the chromatographic band width, and the amount of computer memory needed to handle the matrix [*A*] will be reduced to a minimum, in addition to the calculation time.

In the evaluation of gradient chromatograms where the band widths of the peaks are not only dependent on the properties and efficiency of the column but also on the shape of the gradient, a different approximation must be considered.

When linear gradients are used a constant value of the interval q_{tr} can be used. On the other hand, in stepwise gradients (in fact, those used by the PREOPT package) the band width of the peaks must not necessarily be approximately constant because the gradient shape sometimes include large plateaux. In this instance, band widths of the peaks can be calculated according to the equation[25]

$$w_g = \frac{4\,V_m}{N^{1/2}}(1 + k'_n) \tag{22}$$

where k'_n is the capacity factor that will be obtained in an isocratic elution with a mobile phase composition equal to that of the gradient step in which the peak leaves the column. In such a case, the q_{tr} values must be adapted to the particular shape of the gradient.

From the above discussion, it is evident that the main limitation of informational *OF*s still remains (the lack of applicability to real chromatograms owing to the difficulties of reliable measurement of peak overlapping). However, as the interest in simulation off-line optimization schemes increases, these *OF*s could become a valuable tool.

EXPERIMENTAL ON-LINE OPTIMIZATION *VS.* SIMULATED OFF-LINE OPTIMIZATION. THE ROLE OF THE OBJECTIVE FUNCTION

The automation of experimental optimization in chromatographic separations is a logical consequence of the proposals and studies carried out during the last 15 years, coinciding with a generalized tendency towards the automation of chromatographic techniques, and is at present possible thanks to the low cost of modern microprocessors[26–28].

As Berridge stated[26], there are three basic needs for achieving automatic optimization of chromatographic separations:

(1) the chosen optimization scheme must direct the optimization of the interdependent variables affecting the chromatographic process in an efficient and reliable manner;

(2) each complete chromatogram must be able to be evaluated in terms of the adequate chromatographic parameters (according to Berridge's proposal, the resolution achieved and the time taken by the separation, and it may be necessary to identify each peak by reference to the appropriate standards);

(3) there must exist sufficiently unambiguous bidirectional communication between all the units of the chromatograph and the computer controlling it.

At present, the last condition has been clearly solved, but this is not so with the first two conditions.

A comparison of the different kinds of optimization strategies used in

chromatographic method development is beyond the scope of this paper, but extensive discussion in Schoenmakers' book[1] clearly demonstrates that no actual strategy is free from serious limitations. We agree with Berridge[26] about the power of the simplex method as an optimization strategy, but the multi-modal nature of the response surfaces in chromatography hinders the localization of a global optimum. This is particularly true in gradient and multi-solvent isocratic separations[29]. It is true that this difficulty may be reduced appreciably by carrying out several simplexes in different zones of the response surface, assuming that we would finally obtain convergency in the same optimum, yet this implies lengthening the whole optimization process considerably.

It is also true that by means of an automated instrumental system the time taken in the optimization process is no longer a critical factor, but it is not less true that the experimental realization of such a process implies great expenditure and the equipment will remain out of use (other than for the optimization itself) during this time.

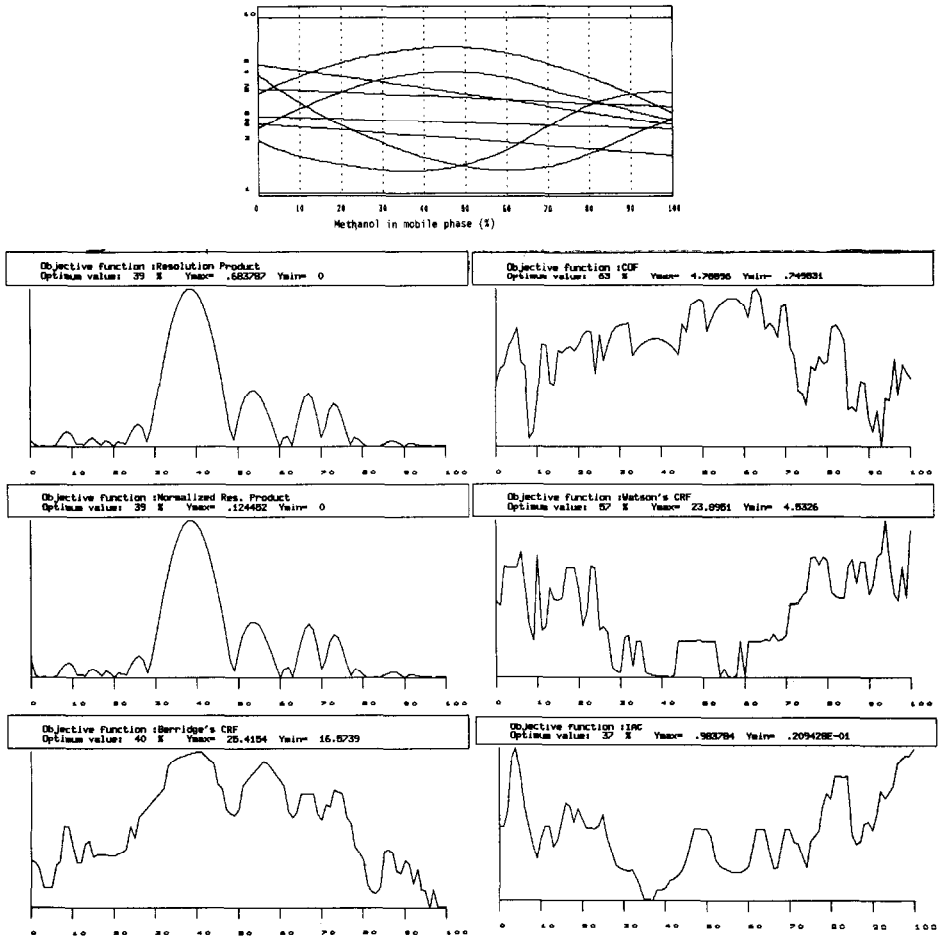

Fig. 4. Response surfaces for six objective functions used to evaluate a simulated chromatographic retention map.

Regarding the second requirement, the *OF* plays the principal role. We have seen before that no single *OF* fulfils all the requirements necessary to consider it as an *OF* of general validity. Moreover, it is necessary to take into account that it is in fact the *OF* which defines the shape of the response surface where the optimization strategy moves. Hence it is pointless to consider both questions in an independent manner. In fact, the multi-modal character of chromatographic response surfaces is the result of two components: (i) the different sets of chromatographic parameter values giving similar chromatograms and (ii) the different chromatograms giving an identical value of the *OF* used. In some instances the second term is the principal component of this sum.

As an example, consider the graphs in Fig. 4, which correspond to response surfaces as shown by several *OF*s. As the chromatograms evaluated are the same and the retention map has been drawn in such a way that time weighting factors do not influence the final *OF* value, the differences in shape between the obtained response surfaces are due only to the particular *OF* considered. In particular, the number of

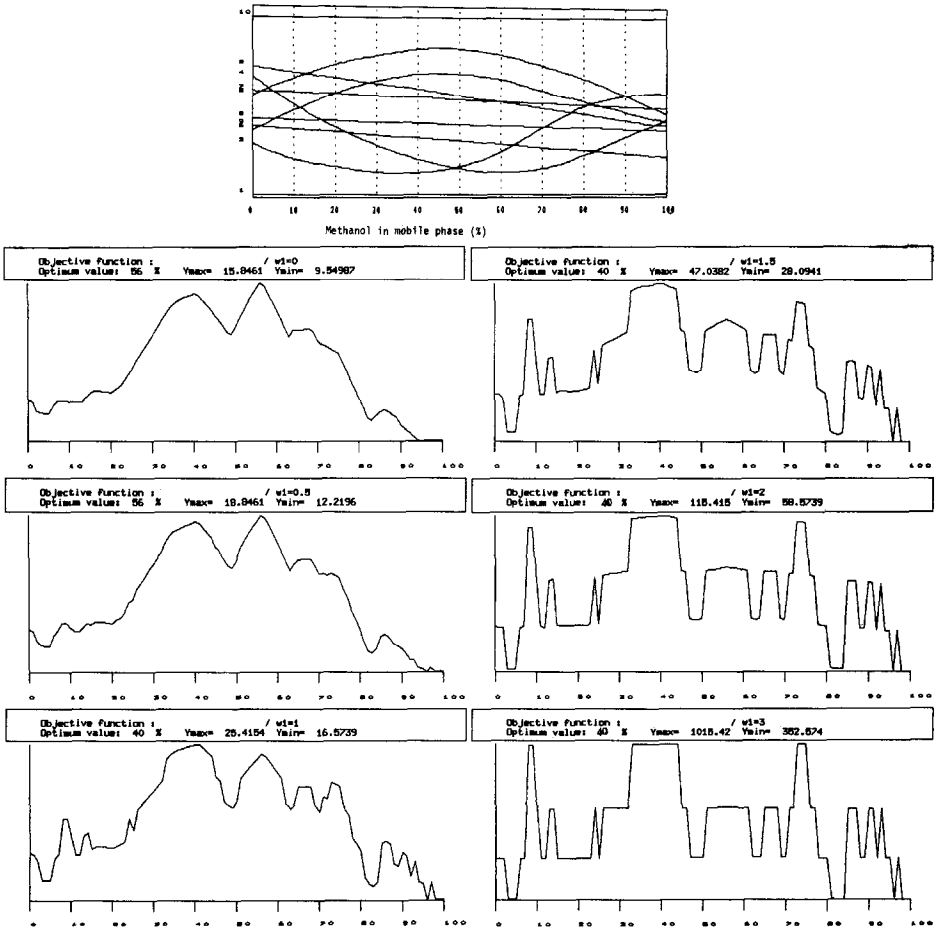

Fig. 5. Influence of the weighting factor w_1 on the performance of Berridge's *CRF*.

maxima (minima in the cases of *CRF* and resolution product *OF*s) varies as a function of the *OF* used.

On the other hand, the weighting factors associated with some *OF*s play an important role in shaping the response surface, thereby producing more or less smoothed surfaces. An example can be seen in the graphs in Fig. 5, corresponding to response surfaces provided by Berridge's *OF* (eqn. 1) as a function of the different values assigned to the w_1 weighting factor. As can be seen, when the w_1 value favours the resolution term in the OF ($w_1 = 0$) a more smoothed response surface is obtained. In contrast, when $w_1 = 2$ (favouring the peak number term) a much less smoothed curve is obtained. As this *OF* was devised to be used in connection with the simplex method strategy, it is evident that the probability of simplex convergence in a local optimum increases on increasing the w_1 value, thus making the simplex task more difficult.

Obviously, the alternative to automated on-line chromatographic optimization is not new. In fact, the first studies on formal chromatographic optimization relied on this approach because instrumental interfacing facilities were not as good as today. These methods (for example, window diagrams[30] or the use of the already mentioned *OF*s) start with sets of experimental data using two approaches: (a) pre-planned experiments, from which to obtain the necessary data in order to use the algorithms efficiently, and (b) sequential experiments, in which the new conditions for the next experiment were obtained from the previous ones after adequate evaluation.

The next step was to simulate the chromatograms as the knowledge of the liquid chromatographic processes develops. This approach has several advantages over experimental optimization[31–34]:

(1) The computer simulation of chromatographic separations avoids most of the experimental work to be done in chromatographic method development and optimization.

(2) Consequently, the cost and time spent in the optimization process are dramatically reduced.

(3) Once the simulation process begins, it can continue in an unattended manner.

(4) Only the computer is blocked during the optimization process (except those computers using multi-task operating systems in which optimization can be carried out in the dead-times of routine analysis) and not the chromatograph, which can be used for other purposes.

Obviously, a number of requirements must be fulfilled by these simulation strategies in order to be of real use:

(1) The chromatogram simulation should be achieved with errors smaller than or equal to experimental chromatogram-to-chromatogram variations. In other words, the simulated chromatograms for a given set of controllable parameters must adequately resemble the experimentally attainable ones (under the same conditions) to be able to apply any evaluation criterion with reliability.

(2) The quality of the simulations must not degrade appreciably during the optimization process.

(3) The final proposal (set of chromatographic conditions) must be sufficiently close to the real global optimum, so that it may be put into practice with a small number of testing and fine-tuning experiments.

The first condition is directly related to the simulation model itself. In the

literature a number of proposals[22,34–46] can be found by means of which the retention time of the peaks in isocratic or gradient elutions can be simulated with errors in the range 0–5%. Most of these methods are based on theoretical relationships which make use of a few experimental gradients to devise equations allowing the prediction of the fundamental parameters for subsequent isocratic separations. Others use isocratic experimental data to find empirical equations to predict the behaviour of the mixture in gradient runs. A comparative study of some of these methods was published by Schoenmakers and Blaffert[42]. In order to be of practical use, these models must simulate the experimental values to within 1% or less, in terms of the capacity factor. The reduction of these errors to acceptable limits has been studied for some of these models[41,43].

The second condition is related not only to the reliability of the simulation model but also to the optimization strategy used to handle the simulated chromatograms. In the case of pre-planned strategies (statistical designs), the main errors are associated with the fit of adequate polynomial functions used to extrapolate the retention time values for conditions outside the region studied experimentally[36–40,42,47]. On the other hand, when the simplex method is used[22], the cumulative character of errors in the sequential process and the lack of linearity in the relationship between retention and solvent composition causes the main errors in the final simulated optimum chromatogram.

The third condition is again related to the *OF* used to conduct the optimization process. Obviously, carrying out the optimization in an experimental or simulated manner has no influence on the efficiency of the *OF* in locating a global optimum or in producing an adequate response surface on which the optimization strategy can move. From the above discussion, it is clear that one of the main problems in order to accomplish a good optimization is the selection of the most adequate *OF*. This decision can be more or less difficult to make depending on the knowledge we have about the particular mixture to be separated, but in the worst case the unique approximation is to check several *OF*s in order to choose the most appropriate one for a specific problem[48]. In this situation, the benefits associated with the simulation off-line strategies are evident because the primary data set is the same irrespective of the *OF* assayed. Thus, with an initial data set we can try as many *OF*s as we want, obtaining a picture of the response surface (in pre-planned strategies) or of the optimum (in sequential strategies) given by each particular *OF*. This process will be very rapid and inexpensive because it is the computer which works and not the chromatograph.

Let us consider an example using the PREOPT package[22] which allows us to use the *CRF* of Watson and Carr[10] (in the modified form published by Debets *et al.*[14]), the *CRF* of Berridge[6] (named here as *CRFM*), the *COF* of Glajch *et al.*[11], the *RP* of Schoenmakers[1] or the *IAC* as developed by us. In the PREOPT these *OF*s can be used successively with the same primary data set (a set of isocratic retention time measurements) to optimize binary gradients of any shape.

Fig. 6a shows the retention map for sixteen phenolic compounds on a μBondapak C_{18} column. Fig. 6b shows the best achievable isocratic separation of these compounds using this column and methanol as the organic modifier (according to the result given by the APTA algorithm[49]). Fig. 6c shows the result for a linear binary gradient from 5 to 50% of methanol in 30 min. This figure corresponds to a typical graphical output in PREOPT runs. This graphical output is composed of the following:

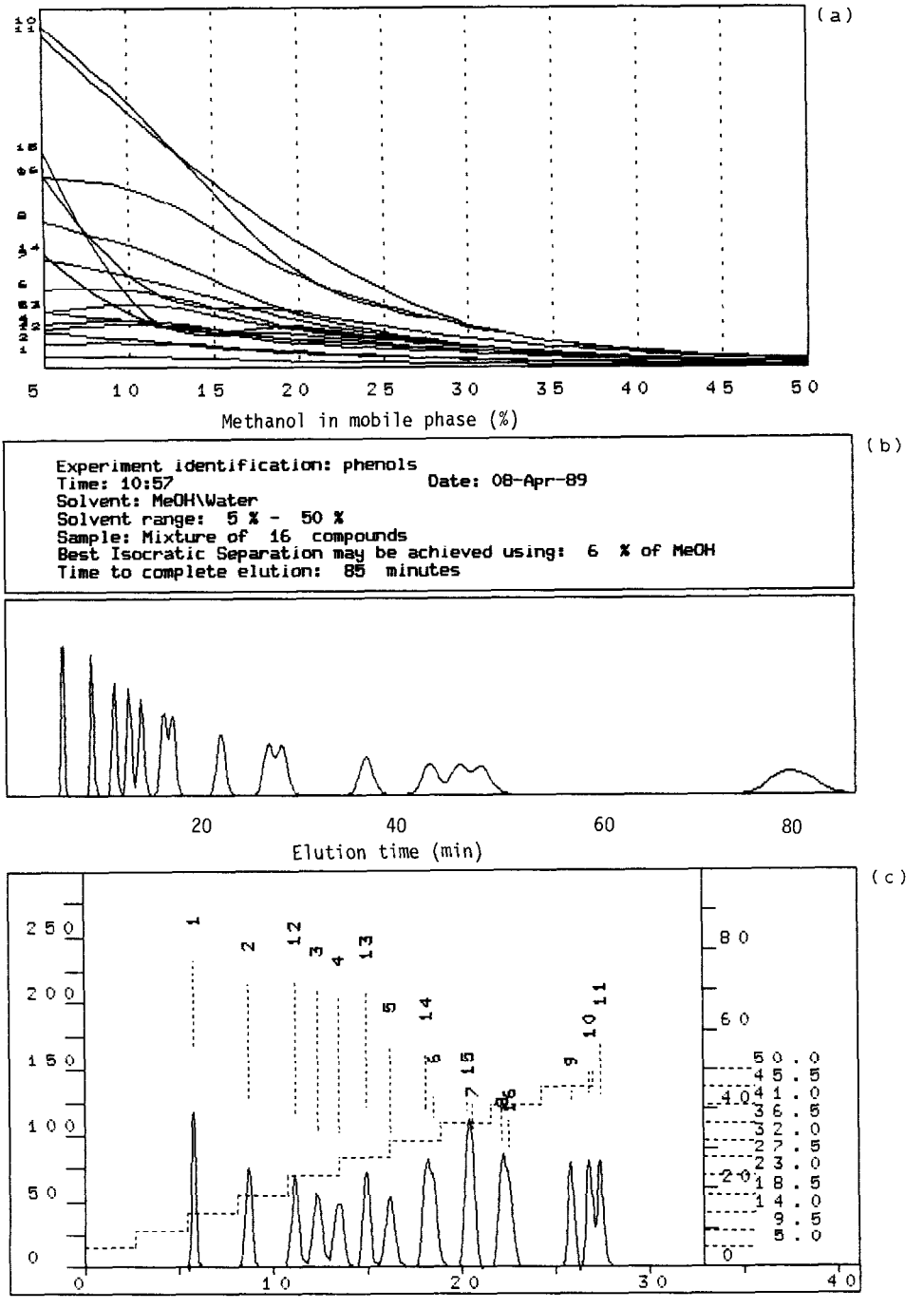

Fig. 6. Retention map, isocratic and linear gradient separations of a mixture of sixteen phenolic compounds. Results of a typical PREOPT run. MeOH = Methanol.

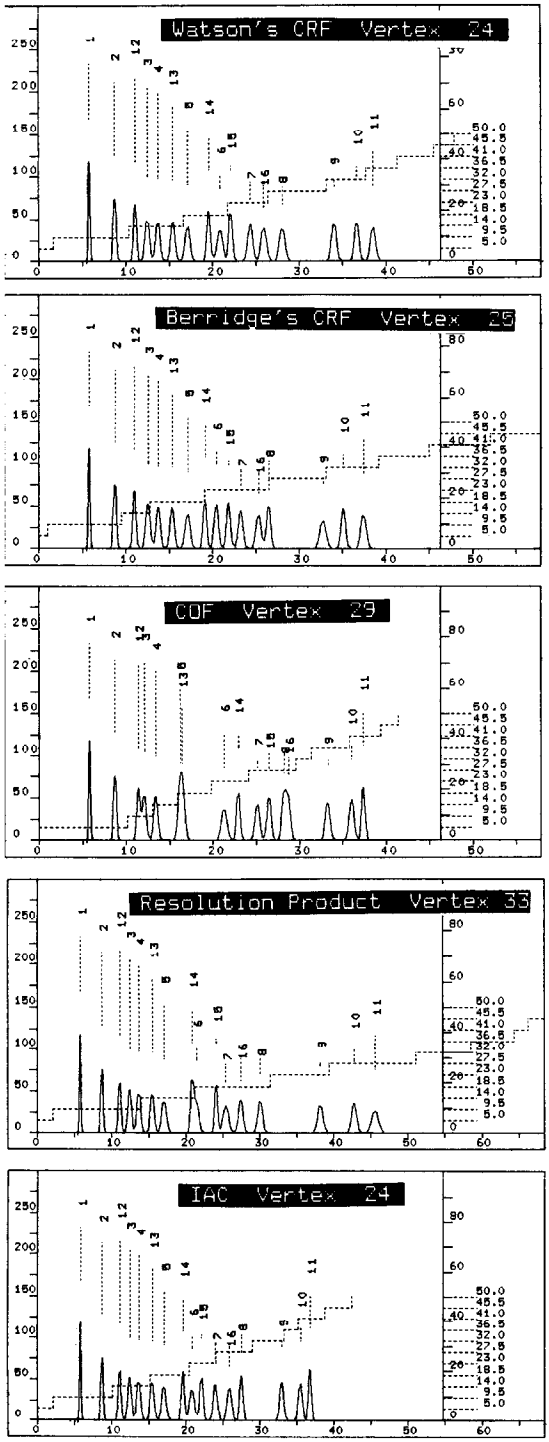

Fig. 7. Optimum separations suggested by the PREOPT package for the mixture of phenolic compounds ii Fig. 6 as a function of the optimization criteria used in the simplex search.

(1) The simulated chromatogram in which the peaks are identified on the top by means of a numerical key corresponding to the order in which the isocratic primary data were introduced into the data base.

(2) The gradient shape using the stepwise mode. The time scale (bottom of the graph) serves to indicate both the retention time of the peaks and the time spent in each gradient stage.

(3) The scale on the right of the graph corresponds to the composition of the mobile phase during the gradient run and consists of a double numerical scale; on the left is the absolute composition (0–100% of the modifier), and the number on the right means the actual gradient used (*i.e.*, the percentage of modifier for each horizontal step).

(4) The scale on the left of the graph is an arbitrary height scale of the peaks.

In the PREOPT output all this information is given also in numerical form by means of tables not shown in Fig. 6.

Fig. 7 shows the optima for the binary (methanol–water) gradient separations of this complex mixture in accordance with the *OF* used in simplex optimization. In all instances the same initial simplex has been used, starting from the linear gradient depicted in Fig. 6c. The simplex type used was the modified simplex method (MSM) of Nelder and Mead[50]. In Fig. 7 the final (optimized) stepwise gradients are also depicted and the methanol percentages corresponding to each step in the gradient are written on the right of each graph. The vertex number corresponding to the optimum is indicated.

As can be seen, very different situations are obtained depending on the particular *OF* used. Some of the chromatograms are considerably better than others, indicating the different abilities of the various *OF*s to conduct the simplex search although the region of the response surface initially explored by the simplex is the same. In this particular example the *CRF*, *CRFM* and *IAC* give better results than the *COF* or *RP* functions, but this conclusion is far from being extrapolatable to other examples. In our experience with several different mixtures, the *COF* function generally gives the poorest results but no reliable conclusion has been obtained regarding which of the remaining functions is the best.

As each complete simulated optimization process takes 15–20 min of unattended operation (except for the *IAC*, where about 90 min are needed), it is possible to carry out a complete screening with the PREOPT in 3–4 h. At the end of this time we draw a conclusion as to which seems to be the best suited *OF* for our particular mixture. As the PREOPT allows the chromatographer to work in a fully interactive way, we can study the globality of the obtained optimum initiating the simplex in other zones of the response surface, changing the simplex type or size, etc., making use of the selected *OF*. In contrast, it is evident that carrying out this process in an experimental way would be disappointing even if a fully automated system were available.

ACKNOWLEDGEMENTS

The development of the *IAC* algorithm was supported by CAICYT of the Spanish Ministry of Education and Science (Project 1189/84). The PREOPT package was developed by one of the authors by means of a research agreement with Millipore–Waters (Milford, MA, U.S.A.).

REFERENCES

1 P. J. Schoenmakers, *Optimization of Chromatographic Selectivity —a Guide to Methods Development*, Elsevier, Amsterdam, 1986.
2 J. C. Berridge, *Techniques for the Automated Optimization of HPLC Separations*, Wiley, Chichester, 1985.
3 S. T. Balke, *Quantitative Column Liquid Chromatography —a Survey of Chemometric Methods*, Elsevier, Amsterdam, 1984.
4 W. Wegscheider, E. P. Lankmayr and M. Otto, *Anal. Chim. Acta*, 150 (1983) 87.
5 P. J. Schoenmakers, *J. Liq. Chromatogr.*, 10 (1987) 1865.
6 J. C. Berridge, *J. Chromatogr.*, 244 (1982) 1.
7 J. C. Berridge, *Microproc. Microsystems*, 7 (1983) 19.
8 P. J. Schoenmakers, A. C. J. H. Drouen, H. A. H. Billiet and L. de Galan, *Chromatographia*, 15 (1982) 688.
9 A. C. J. H. Drouen, P. J. Schoenmakers, H. A. H. Billiet and L. de Galan, *Chromatographia*, 16 (1982) 48.
10 M. W. Watson and P. W. Carr, *Anal. Chem.*, 51 (1979) 1835.
11 J. L. Glajch, J. J. Kirkland, K. M. Squire and J. M. Minor, *J. Chromatogr.*, 199 (1980) 57.
12 D. R. Van Hare and L. B. Rogers, *Anal. Chem.*, 57 (1985) 628.
13 H. J. G. Debets, J. W. Weyland and D. A. Doornbos, *Anal. Chim. Acta*, 150 (1983) 259.
14 H. J. G. Debets, L. B. Bajema and D. A. Doornbos, *Anal. Chim. Acta*, 151 (1983) 131.
15 J. N. Nickel and S. N. Deming, *LC, Mag. Liq. Chromatogr. HPLC*, 1 (1983) 414.
16 D. L. Massart and R. Smits, *Anal. Chem.*, 46 (1974) 283.
17 R. Smits, C. Vanroelen and D. L. Massart, *Fresenius Z. Anal. Chem.*, 273 (1975) 1.
18 H. Kaiser, *Anal. Chem.*, 42 (1970) 24A.
19 V. A. Spencer and L. B. Rogers, *Anal. Chem.*, 52 (1980) 950.
20 K. Eckschlager and V. Stepanek, *Anal. Chem.*, 54 (1982) 1115A.
21 C. Liteanu and I. Rica, *Anal. Chem.*, 51 (1979) 1986.
22 R. Cela, C. G. Barroso, C. Viseras and J. A. Pérez-Bustamante, *Anal. Chim. Acta*, 191 (1986) 283.
23 P. Jandera and J. Churáček, *Gradient Elution in Column Liquid Chromatography —Theory and Practice*, Elsevier, Amsterdam, 1985, Appendices 1 and 2.
24 L. R. Snyder and J. J. Kirkland, *Introduction to Modern Liquid Chromatography*, Wiley-Interscience, New York, 1979, pp. 38–48.
25 P. Jandera and J. Churáček, *Gradient Elution in Column Liquid Chromatography —Theory and Practice*, Elsevier, Amsterdam, 1985, pp. 131–134.
26 J. C. Berridge, *Trends Anal. Chem.*, 3 (1984) 5.
27 J. C. Berridge, *Anal. Chim. Acta*, 191 (1986) 243.
28 J. C. Berridge, *Chem. Br.*, 23 (1987) 1063.
29 F. Dondi, Y. D. Mahie, G. Lodi, P. Reschiglian, C. Pietrogrande, C. Bighi and G. P. Martoni, *Chromatographia*, 23 (1987) 844.
30 R. J. Laub and J. H. Purnell, *J. Chromatogr.*, 161 (1978) 49.
31 L. R. Snyder, J. W. Dolan and M. A. Quarry, *Trends Anal. Chem.*, 6 (1987) 106.
32 M. A. Quarry, R. L. Grob, L. R. Snyder, J. W. Dolan and M. P. Rigney, *J. Chromatogr.*, 384 (1987) 163.
33 L. R. Snyder and J. W. Dolan, *Am. Lab.*, 18 (1986) 37.
34 J. W. Dolan and L. R. Snyder, *LC·GC*, 5 (1987) 970.
35 E. Marego, M. C. Gennaro, C. Baiocchi and P. Bertolo, *Chromatographia*, 25 (1988) 413.
36 J. J. Kirkland and J. L. Glajch, *J. Chromatogr.*, 255 (1983) 27.
37 K. Jinno and K. Kawasaki, *J. Chromatogr.*, 298 (1984) 326.
38 Lu Peichang and Lu Xiaoming, *J. Chromatogr.*, 292 (1984) 169.
39 G. D'Agostino and L. Castagnetta, *J. Chromatogr.*, 338 (1985) 1.
40 G. D'Agostino, M. J. O'Hare, F. Mitchell, T. Salomon and F. Verillon, *Chromatographia*, 25 (1988) 343.
41 L. R. Snyder and M. A. Quarry, *J. Liq. Chromatogr.*, 10 (1987) 1789.
42 P. J. Schoenmakers and T. Blaffert, *J. Chromatogr.*, 384 (1987) 117.
43 M. Quarry, R. L. Grob and L. R. Snyder, *Anal. Chem.*, 58 (1986) 907.
44 K. Jinno and E. Tanigawa, *Chromatographia*, 25 (1988) 613.
45 R. H. Snider, C. F. Moore, E. S. Nylen and K. L. Becker, *BioChromatography*, 3 (1988) 100.
46 Y. Sakamoto, N. Kawakami and T. Sasagawa, *J. Chromatogr.*, 442 (1988) 69.
47 J. L. Glajch, J. J. Kirkland and J. M. Minor, *J. Liq. Chromatogr.*, 10 (1987) 1727.
48 A. K. Smilde, C. H. P. Bruins, D. A. Doornbos and J. Vink, *J. Chromatogr.*, 410 (1987) 1.
49 R. Cela, I. Sanchez and J. A. Pérez-Bustamante, *Intell. Instrum. Comput.*, (1987) 231.
50 J. A. Nelder and R. Mead, *Comput. J.*, 7 (1965) 308.