Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1018 (2003) 169–181
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

# Error analysis and performance of different retention models in the transference of data from/to isocratic/gradient elution

G. Vivó-Truyols, J.R. Torres-Lapasió*, M.C. García-Alvarez-Coque

*Departamento de Química Analítica, Universitat de València, c/Dr. Moliner 50, 46100 Burjassot, Spain*

Received 10 March 2003; received in revised form 14 July 2003; accepted 18 August 2003

**Abstract**

The transferability of retention data among isocratic and gradient RPLC elution modes is studied. For this purpose, 16 β-blockers were chromatographed under both isocratic and gradient elution with acetonitrile–water mobile phases. Taking into account the elution mode where the experimental data come from, and the mode where the retention should be predicted, the following combinations are possible: isocratic predictions from (i) isocratic or (ii) gradient experimental designs; and gradient predictions from (iii) isocratic or (iv) gradient data. Each of these possibilities was checked using three retention models that relate the logarithm of the retention factor: (a) linearly and (b) quadratically with the volume fraction of organic solvent, and (c) linearly with a normalised mobile phase polarity parameter. The study was carried out under two different perspectives: a straightforward examination of the prediction errors and the analysis of the uncertainties derived from the variance–covariance matrix of the fitted models. The best combinations of prediction mode and model were: (i)–(b), (ii)–(c), (iii)–(b), and (iv)–(a) or (c).


*Keywords:* Retention prediction; Retention models; Error analysis; Isocratic elution; Gradient elution; Linear solvent strength theory

## 1. Introduction

High-performance liquid chromatography (HPLC) has become the most applied separation technique for research and routine analysis purposes. Several factors have contributed to this growth. One of them is the possibility of finding the conditions for an optimal separation in an easy and reliable way. The most suited tools to obtain the best separation conditions are, undoubtedly, interpretive optimisations, which substitute successfully the more intuitive—although less efficient—trial-and-error approaches. Several optimisation software packages using interpretive methodologies have been developed. They expedite greatly the analytical method development and assist inexperienced users to set-up separations [1–5].

HPLC separations can be faced under two perspectives, namely isocratic and gradient elution, depending on the retention factor ranges expected for the involved solutes. The former mode has undergone a stronger development, which can be attributed to a less demanding instrumentation and its more simple fundaments. As a consequence, many reports are found in the literature focused on the prediction of the retention

* Corresponding author. Tel.: +34-96-3543003; fax: +34-96-3544436.
*E-mail address:* jrtorres@uv.es (J.R. Torres-Lapasió).

0021-9673/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2003.08.044

in isocratic mode [6–9]. Retention modelling has also been investigated in both normal [10,11] and reversed phase [11,12] gradient chromatography. The accuracy of computer simulation of such predictions has been studied to anticipate and minimise the impact of predictive errors when the linear solvent strength theory is used [13].

The prediction of the retention under gradient conditions is usually done using information coming from gradient experiments. Similarly, runs at a constant mobile phase composition are routinely used to predict the isocratic retention. Crossing the information between both elution modes is also possible: data from gradient experiments can be used to anticipate the retention in isocratic runs [14,15], and vice-versa [16]. However, the quality of the predictions, especially when the information obtained in one elution mode is used to predict the retention in the other, is still a subject of study.

Any optimisation can be divided in two steps, namely modelling of retention (whose accuracy has a decisive influence on resolution measurements), and the optimisation itself. Dividing the process in such steps simplifies the treatment and makes the conclusions more easily interpretable. This report concerns the modelling step and discusses how to process the information gathered in either isocratic or gradient RPLC experimental designs to predict the retention in both elution modes, in order to get a maximal benefit for a subsequent optimisation. The limitations associated to the selected experimental designs using different retention models are studied. The way to overcome them is shown. The quality of the predictions is evaluated by checking the uncertainty of the models.

## 2. Theory

### 2.1. Prediction of the retention

In isocratic mode, the retention behaviour of a given solute in RPLC can be described by establishing a polynomial relationship between the logarithm of the retention factor, $k$, and the volume fraction of organic solvent in the aqueous–organic mobile phase, $\varphi$. This dependence has been proposed to be quadratic [17]:

$$\log k = c_0 + c_1\varphi + c_2\varphi^2 \quad (1)$$

where $c_i$ are regression coefficients with characteristic values for a given solute and column–solvent system. In narrow concentration ranges of organic solvent, a linear dependence can be valid:

$$\log k = c_0 + c_1\varphi = \log k_w - S\varphi \quad (2)$$

$k_w$ being the retention factor using pure water as mobile phase and $S$ a parameter related to the elution strength. A similar two-parameter model, appropriate for wider concentration ranges, substitutes the volume fraction of organic modifier by a normalised polarity parameter, $P_m^N$ [6]:

$$\log k = c_0 + c_1 P_m^N \quad (3)$$

where again $c_i$ are regression coefficients. For acetonitrile–water mixtures, the mobile phase polarity parameter is defined as:

$$P_m^N = 1 - \frac{1.33\varphi}{1 + 0.47\varphi} \quad (4)$$

Eqs. (1)–(3) can be fitted using experimental data from either isocratic or gradient experiments. Isocratic modelling involves a straightforward fitting of the retention data, which can be improved by including weighting factors to compensate the heteroscedasticity introduced by the logarithmic transformation of the original response [18]. A second possibility consists of fitting the data taken from several gradient runs. For this purpose, the retention of a solute eluting under a given gradient program should be expressed as a function of the retention parameters in Eqs. (1)–(3). The fundamental integral equation for gradient elution is:

$$t_0 = \int_0^{t_g - t_0} \frac{dt}{k(\varphi(t))} \quad (5)$$

where $t_0$ is the dead time, $t_g$ the retention time of the solute under gradient conditions, and $k(\varphi(t))$ the equation describing the solute retention factor at the column inlet as a function of time. From this equation, the retention time can be calculated for any gradient, provided that $k(\varphi(t))$ is known. This dependence is established by introducing the programmed gradient, $\varphi = f(t)$, in the retention model, $k = f(\varphi)$ (Eqs. (1)–(3)).

The linear solvent strength theory developed by Snyder and Dolan [19] demonstrates that if a linear dependence between $\log k$ at the column inlet and time

holds (i.e. Eq. (2) is valid and the gradient program is linear), Eq. (5) has the following algebraic solution:

$$t_g = \frac{t_0}{b} \log (2.3 k_0 b + 1) + t_0 + t_D \quad (6)$$

where $b$ is related to the solvent strength, the slope of the gradient program ($\varphi'$), and $t_0$, through:

$$b = S\varphi' t_0 \quad (7)$$

being

$$\varphi' = \frac{\Delta\varphi}{t_G} \quad (8)$$

where $t_G$ is the gradient time and $\Delta\varphi$ the difference between final and starting organic solvent concentrations. In Eq. (6), $t_D$ is the time delay till the gradient reaches the column inlet (dwell time), and $k_0$ the retention factor at the beginning of the gradient:

$$k_0 = 10^{(\log k_w - S\varphi_0)} \quad (9)$$

$\varphi_0$ represents the mobile phase composition when the gradient starts. For linear gradients:

$$\varphi = \varphi_0 + \varphi' t \quad (10)$$

When the solute migrates inside the column during the dwell time (i.e. pre-elution), Eq. (6) must be rewritten as [20,21]:

$$t_g = \frac{t_0}{b} \log [2.3 k_0 b(1 - f) + 1] + t_0 + t_D \quad (11)$$

where $f$ indicates the column fraction migrated by the solute before being affected by the gradient:

$$f = \frac{t_D}{t_0 k_0} \quad (12)$$

The solution of Eq. (5) depends on the retention model. Eq. (11) can only be applied when a linear relationship exists between $\log k$ and $\varphi$ (Eq. (2)). It should be noted that Eq. (5) has no algebraic solution when the retention is described by Eq. (3), since the dependence between $k$ and $\varphi$ is non-linear. Meanwhile, Eq. (5) combined with Eq. (1) has an algebraic solution, although depending on the error function [17]. In this case, an iterative method is required, since the gradient retention time is included in the upper limit of the integral. In this work, numerical integration (see Section 4.1) was used with Eqs. (1) and (3).

Data from isocratic experiments were linearly fitted to obtain the model parameters ($c_i$ in Eqs. (1)–(3)). For gradient experimental designs, an iterative process was needed owing to the non-linear relationship between the retention time and the model parameters. In this work, the quadratically-convergent Powell method [22] was used for these fittings.

The main objective of this work was to examine the prediction of the retention in either isocratic or gradient modes, when data coming from isocratic or gradient experiments are modelled using different equations. Four source–target combinations of data transference were hence considered (i.e. isocratic–isocratic, isocratic–gradient, gradient–isocratic, gradient–gradient). Here, the term "source" will be used to refer the elution mode where the experiments were run to fit the data, and "target", the elution mode where the predictions are performed.

### 2.2. Stepwise numerical integration

As commented, in contrast to the straightforward calculation of retention times under isocratic elution, gradient conditions require frequently numerical integration of Eq. (5). The upper limit contains the unknown variable, $t_g$, which can be found by splitting the integral in small isocratic steps (i.e. increasing gradually the independent variable, $t$):

$$t_0 = \int_0^{t_g - t_0} \frac{dt}{k(t)} = \int_0^{t_1} \frac{dt}{k(t)} + \int_{t_1}^{t_2} \frac{dt}{k(t)} + \cdots + \int_{t_{i-1}}^{t_i} \frac{dt}{k(t)} + \int_{t_i}^{t_{i+1}} \frac{dt}{k(t)} \quad (13)$$

$k(t)$ can be assumed to be constant in each of these steps. In this case, $t_0$ can be approximated to:

$$t_0 \approx \frac{t_1}{k_{0,1}} + \frac{t_2 - t_1}{k_{1,2}} + \cdots + \frac{t_i - t_{i-1}}{k_{i-1,i}} + \frac{t_{i+1} - t_i}{k_{i,i+1}} = I_{0,i} + I_{i,i+1} = I_{0,i+1} \quad (14)$$

with:

$$k_{i,i+1} = \frac{k(t_i) + k(t_{i+1})}{2} \quad (15)$$

The approximate cumulative integral $I_{0,i}$ is monitored up to fulfil the condition $I_{0,i} < t_0 < I_{0,i+1}$. At this point, $t_g$ can be easily obtained as:

$$t_g = t_0 + t_i + (t_0 - I_{0,i}) k_{i,i+1} \quad (16)$$

The influence of the step size $(t_{i+1} - t_i)$ on the precision of the predictions will be discussed in Section 4.1.

### 2.3. Confidence intervals

Confidence intervals are a convenient way to assess the quality of retention time predictions with regard to the experimental factors (e.g. solvent concentration or gradient slope). These intervals are computed in two steps. The first one is the calculation of the variance–covariance matrix of the regression parameters in the retention model, which concerns the uncertainties introduced by the source data. The second step is the propagation of these uncertainties in the prediction of retention times, which concerns the target mode.

Besides time measurement errors, for a given solute, the variance–covariance matrix depends on the retention model and the experimental design of the source data. In the absence of lack of fit (i.e. the model describes correctly the retention), this matrix is given by [23]:

$$\mathbf{V} = S^2_{\text{e,source}}[(\mathbf{J}^{\text{T}}_{\text{source}}{}^{\text{source}}\mathbf{J}^{\text{source}}_{\text{source}})^{-1}] \tag{17}$$

where $S^2_{\text{e,source}}$ is the pure experimental error of the source data, which for a design including $n$ points to fit a model with $m$ parameters can be approximated to:

$$S_{\text{e,source}} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{t}_{R_i^{\text{source}}} - t_{R_i^{\text{source}}})|_{\mathbf{b}}}{n-m}} \tag{18}$$

$\hat{t}_{R_i^{\text{source}}}$ and $t_{R_i^{\text{source}}}$ are predicted and experimental retention times, respectively, at the $i$th point of the experimental design for the source elution mode. The calculated retention times are evaluated using the parameter set $\mathbf{b} = (b_1, b_2, \ldots, b_m)$, which is obtained by fitting the source experimental data to the considered retention model (i.e. $\mathbf{b}^{\text{iso}}$ for the parameters fitted from isocratic data and $\mathbf{b}^{\text{grd}}$ from gradient data).

The Jacobian matrix in Eq. (17), $\mathbf{J}^{\text{source}}_{\text{source}}$, measures the sensitivity of time predictions in the source mode to errors in the regressed parameters. The $ij$th element of the $n \times m$ sized $\mathbf{J}^{\text{source}}_{\text{source}}$ matrix is defined as:

$$j^{\text{source}}_{\text{source},ij} = \left.\frac{\partial \hat{t}_{R_i^{\text{source}}}}{\partial b_j}\right|_{\mathbf{b}} \tag{19}$$

Each $j^{\text{source}}_{\text{source},ij}$ term in the Jacobian matrix is the derivative of the $i$th predicted retention time in the source mode, when the $j$th regression parameter in $\mathbf{b}$ is varied.

Eq. (17) can be used to calculate the uncertainties of the predictions performed in the target elution mode. The standard deviations of a set of experiments will be given by:

$$\begin{aligned}\mathbf{s}_{\text{target}} &= \sqrt{\mathbf{J}^{\text{source}}_{\text{target}}\mathbf{V}\,\mathbf{J}^{\text{source}^{\text{T}}}_{\text{target}}} \\ &= S_{\text{e,source}}\sqrt{\mathbf{J}^{\text{source}}_{\text{target}}(\mathbf{J}^{\text{source}^{\text{T}}}_{\text{source}}\mathbf{J}^{\text{source})^{-1}}_{\text{source}}\mathbf{J}^{\text{source}^{\text{T}}}_{\text{target}}}\end{aligned} \tag{20}$$

where $\mathbf{J}^{\text{source}}_{\text{target}}$ is the Jacobian matrix for the target mode. The size of this matrix, $q \times m$, is given by the number of experiments ($q$) that are being predicted and the number of fitted parameters ($m$) in the source elution mode. The $ij$th element of $\mathbf{J}^{\text{source}}_{\text{target}}$ is defined as:

$$j^{\text{source}}_{\text{target},ij} = \left.\frac{\partial \hat{t}_{R_i^{\text{target}}}}{\partial b_j}\right|_{\mathbf{b}} \tag{21}$$

Note that here, the predicted retention times are referred to the target mode (in contrast with Eq. (19)), but the set of parameters obtained in the source mode ($b$) is used again. Consequently, four source–target Jacobians are possible: $\mathbf{J}^{\text{iso}}_{\text{iso}}$, $\mathbf{J}^{\text{iso}}_{\text{grd}}$, $\mathbf{J}^{\text{grd}}_{\text{iso}}$ and $\mathbf{J}^{\text{grd}}_{\text{grd}}$. The subindex $i$ in Eq. (21) denotes a given experimental condition for which the retention time is predicted (e.g. concentration of organic solvent in isocratic elution, or gradient program in gradient elution). The confidence limits are calculated from the standard deviation considering the $t$-value for $n - m$ degrees of freedom.

## 3. Experimental

### 3.1. Reagents

Sixteen β-blockers were studied: acebutolol (Italfármaco, Alcobendas, Madrid, Spain), alprenolol, pindolol, sotalol (Sigma, St. Louis, MO, USA), atenolol (Zeneca Farma, Madrid), bisoprolol, propranolol, practolol (ICI-Farma, Madrid), carteolol (Miquel-Otsuka, Barcelona, Spain), celiprolol (Rhône-Poulenc Rorer, Alcorcón, Madrid), esmolol (Du Pont-De Nemours, Le Grand Saconnex, Switzerland), labetalol (Glaxo, Tres Cantos, Madrid), metoprolol, oxprenolol

(Ciba-Geigy, Barcelona), nadolol (Squibb, Esplugues de Llobregat, Barcelona), and timolol (Merck, Sharp & Dohme, Madrid). Acebutolol, atenolol, carteolol, celiprolol, labetalol, metoprolol, nadolol, oxprenolol, propranolol, and timolol, were kindly donated by the cited pharmaceutical laboratories. The drugs were dissolved in a small amount of methanol and diluted with water. The concentration of the stock and injected solutions were 100 and 10 μg/ml, respectively. These solutions remained stable for at least 2 months at 4 °C.

Mobile phases were prepared with acetonitrile (Scharlab, Barcelona). The pH was buffered with di-sodium hydrogen phosphate and hydrochloric acid (Panreac, Barcelona) at pH 3 to improve the peak shape of the basic β-blockers. The mobile phases and drug solutions to be injected were vacuum filtered through 0.45 μm Nylon membranes (Micron Separations, Westboro, MA, USA). Nanopure water (Barnstead, Sybron, Boston, MA, USA) was used throughout. Acetone (Guinama, Barcelona) was used to measure the dwell time.

### 3.2. Apparatus

An Agilent (Model HP 1100, Palo Alto, CA, USA) chromatograph, equipped with a quaternary pump, a UV-Vis detector and an autosampler, was used. A PC computer was connected to the chromatograph through an integrator (Model HP 3396A). Signal acquisition was made with the PEAK-96 software (Hewlett Packard, Avondale, PA, USA). An XTerra MS C18 column (150 mm × 4.6 mm i.d., 5 μm particle size) and a guard column of similar characteristics (20 mm × 3.0 mm i.d., 5 μm particle size; Waters, Milford, MA, USA) were used. The detection wavelength was 225 nm for all β-blockers, except timolol, for which it was 300 nm. The flow-rate was 1.0 ml/min, and the injection volume, 20 μl. The whole study was carried out at room temperature. Duplicate injections of each chromatogram were made.

The dead time was measured as the first base-line deviation, and the dwell time as indicated in [24], by running a blank gradient where acetone was increased from 0 to 1% in 20 min. For this determination, the times at the beginning and end of the steep increase were taken. The signal was monitored at 280 nm. Home-built in routines, written in MATLAB 6.5 (The Mathworks, Natick, MA, USA), were developed for data treatment.

## 4. Results and discussion

### 4.1. Precision in the numerical integration

The prediction of retention times under gradient elution requires solving Eq. (5), either algebraic or numerically (Eqs. (13)–(16)). Theoretically, the precision in $t_g$ can be increased as much as wished, by just decreasing the time increment ($t_{i+1} - t_i$) in the split integrals (Eq. (13)). The higher the desired precision, the longer the computation time. Unfortunately, there is a limitation associated with the pumping system: real gradients are generated by the instrument by approximating changes in composition to steps. This makes useless the effort of obtaining a precision level in $t_g$ greater than that allowed by such an instrumental threshold. In our equipment, the minimal variation in the organic solvent content is $d\varphi = 0.1\%$. The time increments used in the numerical integration were fixed taking into account this value.

For gradients involving moderate contents of organic solvent, the stepwise growth in concentration is performed only once the programmed variation reaches the critical level, $d\varphi$. This introduces a delay between programmed and actual gradients, which makes the solvent concentration equal or systematically smaller than the programmed one. When Eq. (5) is numerically solved, this delay can be corrected by modifying the programmed gradient to mimic the actual stepwise gradient. In case of algebraic solution (linear solvent strength theory), this delay, which depends on the gradient slope, can be compensated by the addition of a new term to the dwell time:

$$t'_D = t_D + \frac{d\varphi}{2\varphi'} \quad (22)$$

The dwell time, $t_D$, was calculated as commented in Section 3 by programming a $\Delta\varphi$ enough to make the additional delay in Eq. (22) negligible.

### 4.2. Accuracy in the prediction of retention times

Two experimental designs were carried out, one of them in isocratic and the other in gradient mode.

The ranges of acetonitrile concentration in the mobile phases were selected to avoid retention times close to the void volume or above 60 min. In the isocratic case, chromatographic parameters were obtained from six mobile phases containing 5, 10, 15, 20, 25 and 30% (v/v) acetonitrile at pH 3. Due to the range of polarities of the solute set (octanol–water partition coefficient, $\log P_{o/w} = 1$–3), measurement of retention times in all mobile phases of the design was not feasible or convenient. In gradient mode, the data were obtained from four gradient runs, where the modifier content was increased from 5 to 30% acetonitrile in $t_G = 20$, 30, 40 and 50 min.

As commented, regarding the source of experimental data and the elution mode in which the prediction should be made (i.e. target mode), four possibilities can be considered: prediction of isocratic retention times from either isocratic or gradient data, and prediction of gradient retention times from either isocratic or gradient data. In all cases, Eqs. (1)–(3) were fitted using the procedure outlined in Section 2.1. For each design (isocratic or gradient), all the available experimental points were considered. Since the degrees of freedom were greater than zero (except when modelling the retention of solutes 1–3 using Eq. (1) with the isocratic design), the calculation of confidence intervals in predictions (see Section 2.3) was feasible.

Fig. 1 depicts the errors obtained when the retention data were collected from an isocratic experimental design, using the three mentioned equations. In the prediction of isocratic retention times (Fig. 1(a)–(c)), only Eq. (1)—which contains three parameters—was able to model properly wide ranges of organic solvent concentration (in the example, 5–30%). On the other hand, between the two-parameter equations studied in

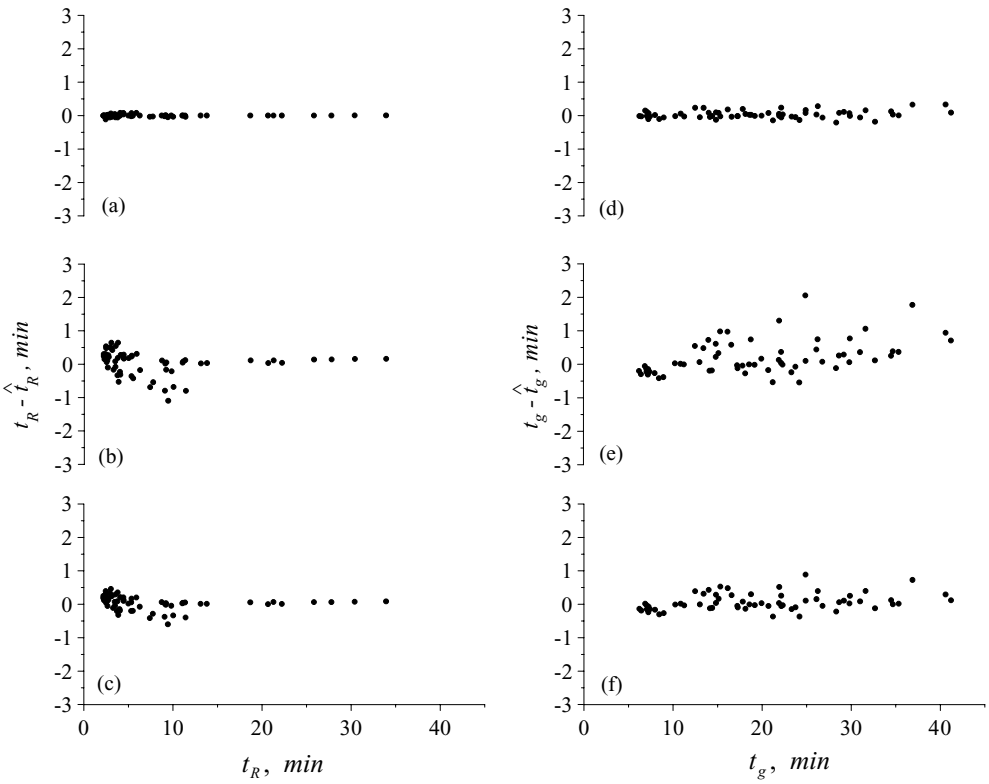


Fig. 1. Accuracy in the prediction of retention data for the set of 16 β-blockers, using measurements from the isocratic experimental design and (a and d) Eq. (1); (b and e) Eq. (2); and (c and f) Eq. (3). Retention data were predicted in (a)–(c) isocratic and (d)–(f) gradient elution modes. Data for all compounds measured in all mobile phases of the experimental designs are plotted altogether.

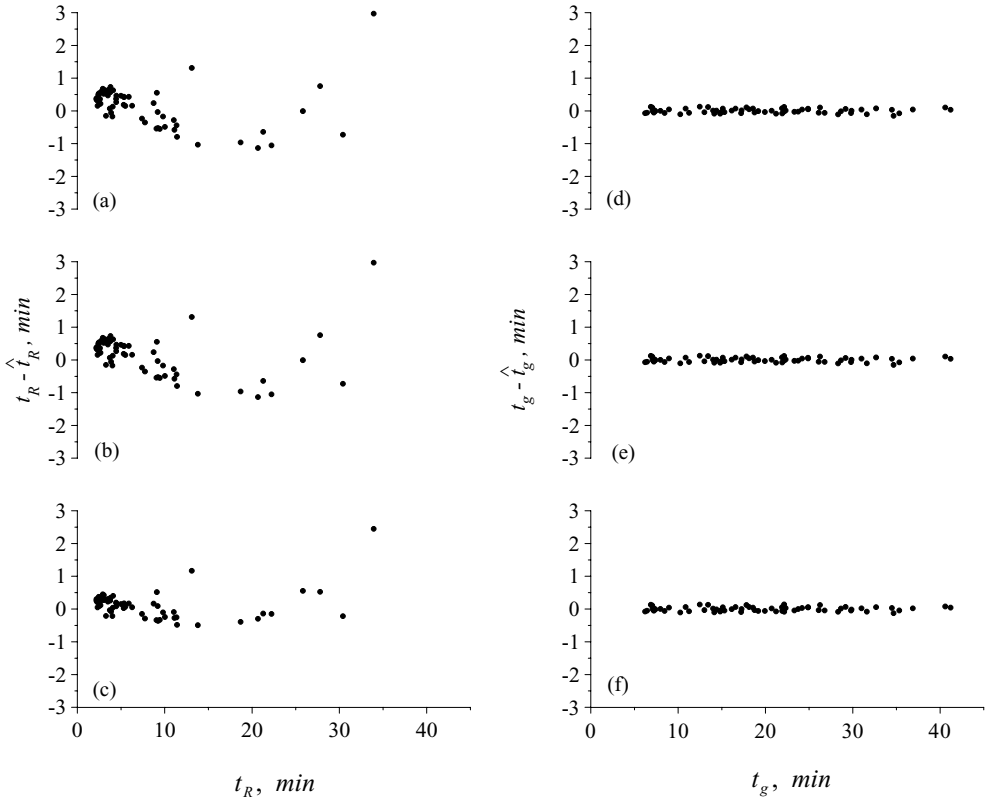


Fig. 2. Accuracy in the prediction of retention data for the set of 16 β-blockers, using measurements from the gradient experimental design. See Fig. 1 for details.

this work (Eqs. (2) and (3)), the latter is preferable, although some bias remains (compare Fig. 1(b) and (c)). Errors obtained in the isocratic fitting are propagated in the prediction of gradient retention. Consequently, only Eq. (1) can be expected to yield unbiased predictions (Fig. 1(d)), whereas Eqs. (2) and (3) yield large errors (Fig. 1(e) and (f)). As in the isocratic case, Eq. (3) showed better accuracy than Eq. (2).

The comparison of the scattering in Fig. 1(a) and (d), where the lack of fit can be considered negligible, allows the estimation of the magnitude of the errors introduced by the transference of data from isocratic to gradient elution. Since the mean error is 0.02 and 0.09 min (maximal errors are 0.12 and 0.33 min) for isocratic and gradient predictions, respectively, the mean uncertainty introduced by the transference of data between both elution modes is 0.07 min. It should be noted that the mean and maximal errors are 0.28 and 1.10 min for Eq. (2) and 0.16 and 0.60 min for Eq. (3) (isocratic predictions), and 0.38 and 2.06 min for Eq. (2) and 0.19 and 0.89 min for Eq. (3) (gradient predictions).

Similar accuracy is observed in the application of different retention models for the prediction of gradient retention times from gradient experimental data (compare Fig. 2(d)–(f)), with mean and maximal errors of 0.06 and 0.15 min, respectively. Such result contrasts remarkably with that observed for the same equations in the prediction of isocratic retention times from isocratic experiments (Fig. 1(a)–(c)). This can be understood by considering the main difference among Eqs. (1)–(3): their capability of modelling the retention in different solvent concentration ranges. Thus, going from Eq. (2) to Eq. (3) to Eq. (1), the solvent domain that can be accurately predicted becomes wider. The information about retention needed

to predict gradient retention times concerns narrower solvent domains than in the isocratic case.

The same problem can be outlined inversely: a gradient experimental design provides information about retention in narrower solvent domains in comparison to the isocratic design. This justifies the poor results yielded when gradient data are used to predict isocratic retention (see Fig. 2(a)–(c)): the experimental domains in isocratic predictions are usually larger than gradient domains, which forces extrapolations. Note that Fig. 2(a) and (b) (and also Fig. 2(d) and (e)) are virtually identical, since the quadratic term in Eq. (1) is non-significant in such fittings. Among the three models, Eq. (3) exhibits the best behaviour in extrapolations (gradient-to-isocratic), and yields, therefore, the best predictions: mean and maximal errors are 0.50 and 2.97 min for both Eqs. (1) and (2), and 0.29 and 2.45 min for Eq. (3).

It is argued that errors in $t_R$ have the same direction for adjacent peaks. Therefore, their impact in the optimisation should be smaller than what is suggested by the errors in $t_R$ for single compounds. In order to check this, errors in $\Delta t_R = t_{R,i+1} - t_{R,i}$ were computed. Fig. 3 shows the deviations between experimental and predicted differences in retention times ($\Delta t_R - \Delta \hat{t}_R$), for all adjacent peaks in the mixture of 16 β-blockers, considering Eq. (2) and the four source–target combinations. As can be observed, the errors in $\Delta t_R$ are similar to those in Figs. 1 and 2. This shows that the errors in retention for adjacent peaks do not cancel each other, and evidences the importance of an accurate modelling. Note that Eq. (2) was the most problematic; for this reason, it was selected for making this comparison.

### 4.3. Confidence intervals and informative range of modifier concentration

A deeper knowledge about the quality of the predictions can be obtained by examining the confidence intervals. The uncertainties in the predictions were calculated as indicated in Section 2.3 for three representative solutes, which were selected attending to their retention: sotalol (fast elution), acebutolol (intermediate), and bisoprolol (slow) Figs. 4–6 depict the isocratic (left plot) and gradient (right plot) predictions, using both isocratic (solid lines) and gradient (dashed lines) source data, together with the 95% confidence limits, for the three models under study (Eqs. (1)–(3)). For sotalol, acebutolol and bisoprolol,

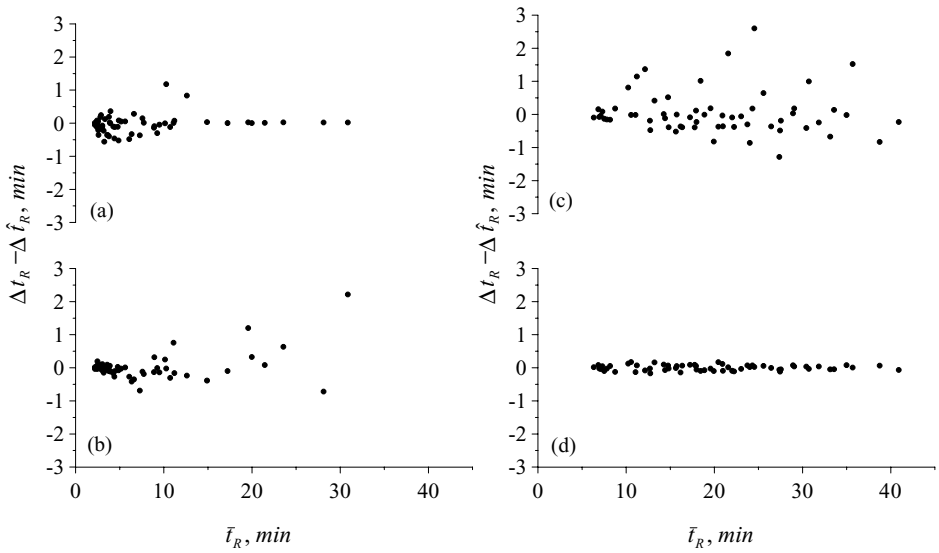


Fig. 3. Errors in the prediction of the difference between retention times for adjacent peaks for the 16 β-blockers against the mean experimental retention time for the peak pair. Eq. (2) was fitted using (a and c) isocratic and (b and d) gradient experimental data, and predictions were made in (a and b) isocratic and (c and d) gradient elution modes.

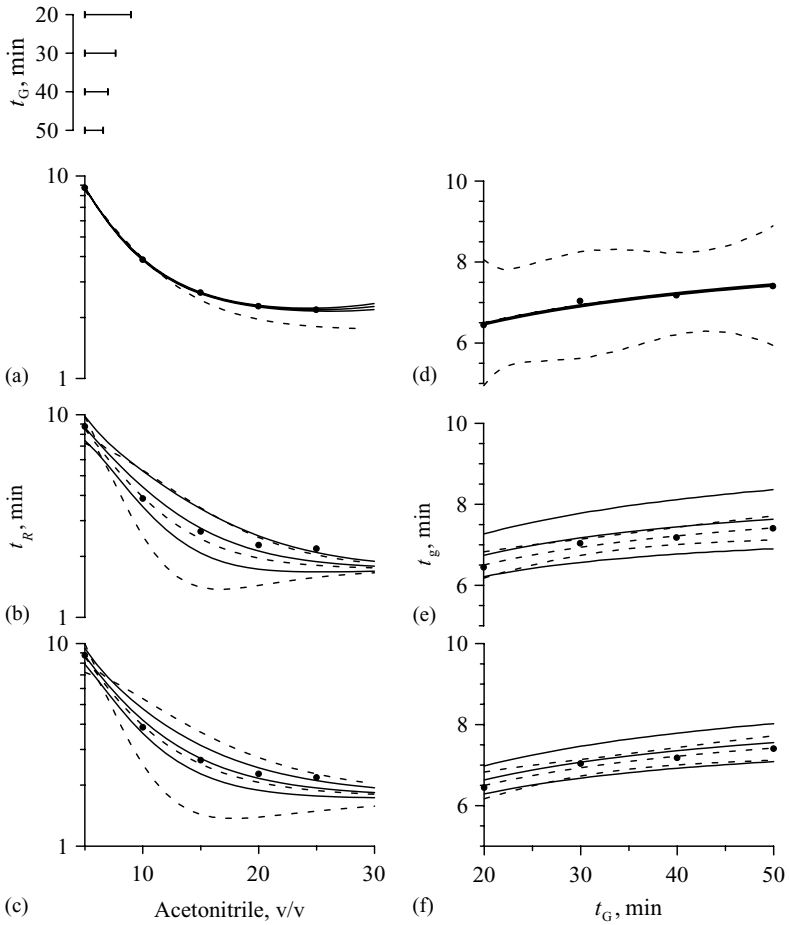


Fig. 4. Prediction of: (a–c) isocratic and (d–f) gradient retention for sotalol using (a and d) Eq. (1), (b and e) Eq. (2), and (c and f) Eq. (3). The 95% confidence limits are given. Source data: (solid line) isocratic and (dashed line) gradient. Experimental data are overlaid (dots). The top of the figure shows acetonitrile concentration domains (*x*-axis scale in the plots below) experienced by the solute during its migration inside the column, for each of the tested gradient programs. The isocratic confidence intervals obtained using gradients and Eq. (1) (a) exceeded the axis scale and were not plotted.

only five (5–25% acetonitrile), five (10−30%) and four (15−30%) isocratic measurements were made, respectively. Outside these ranges the retention times were unpractical.

Solutes chromatographed under gradient elution seldom experience the whole domain of modifier concentration scanned in the gradient program. This means that they usually leave the column before the completion of the gradient, and thus, only a fraction of the programmed solvent concentration range affects their migration. The top of Figs. 4–6 shows the concentration ranges experienced by the selected solutes during their migration inside the column, for each of the four tested gradient programs. These diagrams indicate the actual concentration ranges from which the gradient experiments are able to extract information. These ranges will be called “informative ranges of modifier concentration”. As will be seen, their knowledge allows a more proper comparison of the experimental designs carried out in both isocratic and gradient elution modes.

An adequate interpretation of the confidence intervals requires considering the different factors that in-

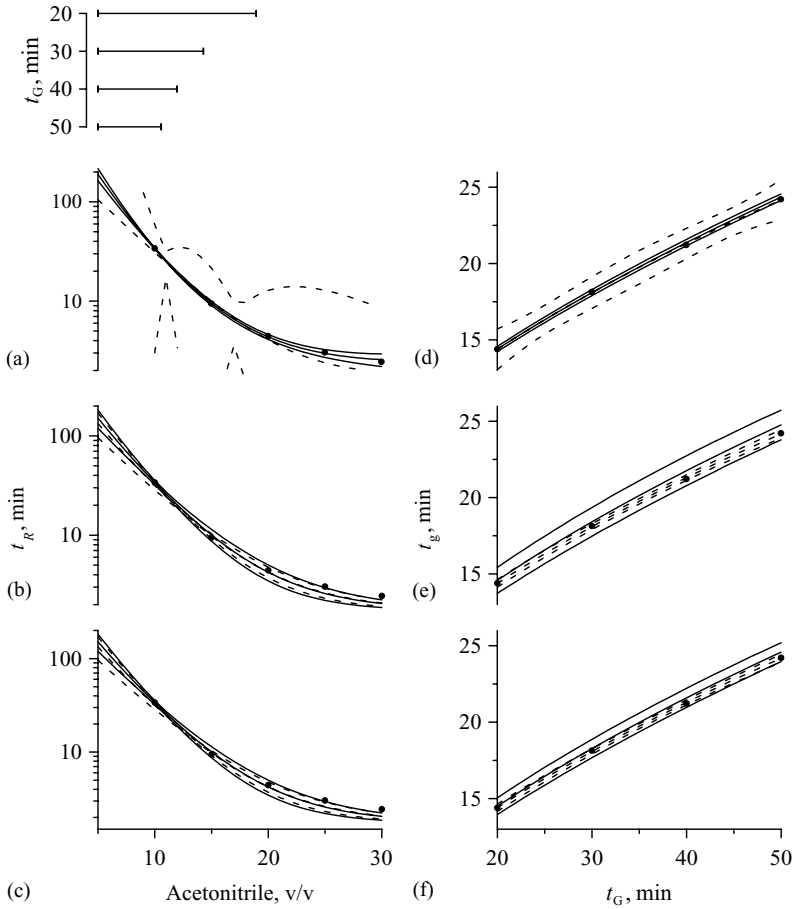


Fig. 5. Prediction of: (a–c) isocratic and (d–f) gradient retention for atenolol, showing the 95% confidence limits. See Fig. 4 for details.

fluence the uncertainties:

(i) The lack of fit.
(ii) The overfitting produced by insufficient degrees of freedom.
(iii) The extrapolation of the retention behaviour in certain concentration regions.
(iv) The existence of modifier concentration domains with negligible migration.
(v) The exponential dependence of the retention factor with solvent concentration.
(vi) The degree of the polynomial retention model (linear or quadratic).

The lack of fit (i) and the overfitting (ii) increase the pure experimental error, $S_{e,source}$ (Eq. (18)), which broadens the confidence interval in the full experimental domain (Eq. (20)). Wider confidence limits are also obtained when extrapolations (iii) are performed. The examination of the modifier concentration ranges for which the retention behaviour is extrapolated is not so straightforward in gradient mode as in the isocratic case. Since solutes usually leave the column before the completion of the gradient, a certain domain of modifier concentration remains unstudied, and the prediction of retention in this region should be extrapolated.

Domains of negligible migration (iv) are of concern when predictions are made using gradient experiments as source data. During a gradient, the solute experiences a progressive increment in the migration rate, and the gradient retention time will mainly depend on the solute behaviour at those mobile phase compositions where this rate is significant. This means that the

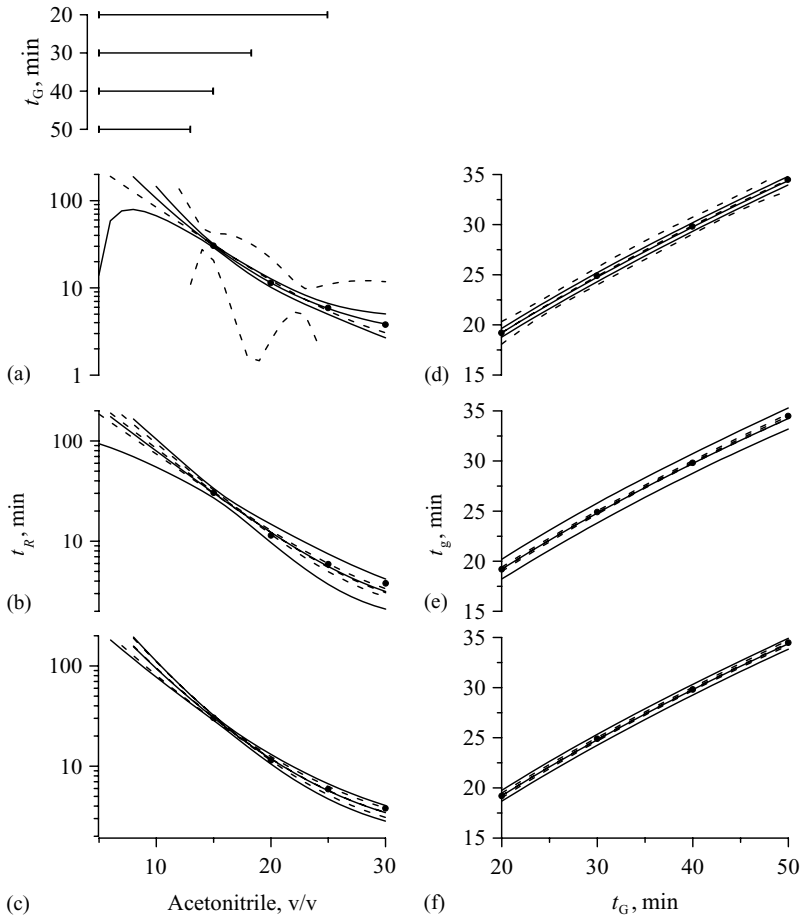


Fig. 6. Prediction of: (a–c) isocratic and (d–f) gradient retention for bisoprolol, showing the 95% confidence limits. See Fig. 4 for details.

observed retention time is insensitive to modifier concentrations of relatively low elution strength, which inflates the confidence limits in such region.

Another factor that influences the confidence interval pattern is the exponential dependence of the retention model (v). At low modifier concentration, the absolute errors are larger due to the higher retention. The use of weighted regression according to [9] certainly diminishes this effect, but even in this case, the trend of the prediction uncertainty remains slightly distorted. Finally, the degree of the polynomial model (vi) determines the number of waists observed in the confidence intervals: one and two waists for linear and quadratic models, respectively, as can be seen in Figs. 4(a) and (d), 5(a) and (d), and 6(a) and (d) (note that the confidence interval is out of scale in Fig. 4(a), being almost linear in Figs. 5(d) and 6(d)).

#### 4.3.1. Isocratic predictions

The factors commented in the previous section are shown up in Figs. 4–6. First, the case of the fastest solute, sotalol, will be discussed (Fig. 4). Several differences exist in the performance of the prediction of isocratic retention, depending on the elution mode of the source data (Fig. 4(a)–(c)). Gradients (dashed lines) yield acceptable predictions for the three retention models only in narrow acetonitrile concentration ranges, which are located approximately between 5 and 10% acetonitrile. Outside this range, the predictions are biased. The upper limit of the informative

solvent concentration range is lower than the composition of the mobile phase at which the solute leaves the column with the steepest gradient. For sotalol, this composition is 9% acetonitrile ($t_G = 20$ min). On the other hand, the lower limit of the informative range is larger than the solvent concentrations at which negligible migration of the solute is observed. As commented in the previous section, the uncertainty is minimal in this informative range, which is detected as a waist in the confidence interval (see dashed lines in Fig. 4(b) and (c) for approximately 6–8% acetonitrile). Due to the extrapolation effect, confidence intervals broaden outside this range, but above 15% acetonitrile they become narrower (see again Fig. 4(b) and (c)), because the exponential decay of the retention factor compensates the extrapolation effect (factor (v)). Finally, when Eqs. (2) and (3) are used for the predictions, the lack of fit makes the experimental values to lie outside the confidence intervals for fast mobile phases.

Isocratic predictions from isocratic experimental designs (solid lines in Fig. 4(a)–(c)) take into account a rather wide range of acetonitrile concentrations. Therefore, only Eq. (1) is valid (Fig. 4(a)), whereas the two-parameter models yield a certain lack of fit (Fig. 4(b) and (c)). This is more evident with Eq. (2).

#### 4.3.2. Gradient predictions

Predictions performed from isocratic designs (solid lines in Fig. 4(d) and (f)) yield confidence intervals usually similar to those achieved for linear regressions, since the dependence of $t_g$ versus $t_G$ is almost linear. In this case, only the informative solvent concentration range participates in the prediction of gradient retention and this range is inside the limits of the isocratic design. Therefore, the uncertainty is nearly independent of $t_G$. Gradients are affected by several sources of error (e.g. pumping instability and dwell time error) that are not considered in the $S_{e,source}$ value used in Eq. (20). This gives rise to an underestimation of the actual errors when Eq. (1) is used. Isocratic retention is usually better predicted (compare Fig. 1(a) and (d)), which certainly decreases $S_{e,source}$. Due to this underestimation, the experimental points lie outside the confidence intervals of the gradient predictions (Fig. 4(d), solid line). In contrast, the lack of fit introduced when the data are modelled using Eqs. (2) and (3) increases $S_{e,source}$, which hides this underestimation. The final result is a larger uncertainty (Fig. 4(e) and (f)). With these two models, the bias produced inside the informative concentration range is translated into systematic errors of the same sign in gradient predictions.

The use of gradient data to fit Eq. (1) gives rise to overfitting, which is translated into an abnormally wide confidence interval (Fig. 4(d), dashed lines). Since Eqs. (2) and (3) do not overfit the data, the confidence intervals agree with the observed scattering (Fig. 4(e) and (f)).

#### 4.3.3. Slower solutes

When going from faster to slower solutes, the informative range of acetonitrile concentration is shifted towards higher values. Thus, the minimum of the confidence interval is located in the region between 11 and 17% acetonitrile for acebutolol, and between 15 and 22% acetonitrile for bisoprolol (Figs. 5 and 6). The concentration range of solvent that affects the solute during the gradient is wider for slower solutes (see top diagrams in the figures). For these solutes, there is a wider concentration domain of negligible migration, and therefore, a larger uncertainty in the predictions (see confidence intervals in Figs. 5(a)–(c) and 6(a)–(c) close to 5% acetonitrile). Note that the scale in the plots is logarithmic.

Another effect to consider is the high relative values of $k_0$ for slower solutes. This allows Eq. (6) to be simplified [15]:

$$t_g \approx \frac{t_0}{b} \log(2.3 k_0 b) + t_0 + t_D \tag{23}$$

This simplification yields errors smaller than 4% for solutes with $k_0 > 45$.

## 5. Conclusions

When facing an unknown separation problem, the chromatographer often lacks of guidelines to decide a priori if isocratic elution will give enough separation in a reasonable analysis time or a gradient will be required. Fast gradients can be useful tools in the first steps of method development to get a quick look about the retention and resolution behaviour, in order to minimise the experimental effort. The final aim in such cases is to determine whether an isocratic

separation would be sufficient. Both situations may require the prediction of retention in an elution mode different from that one in which the data were collected. This implies the transference of the modelled data between different elution modes. In this work, we examine the quality of predictions of retention when the information is crossed using three different models.

Gradient experiments give accurate information about retention only in narrow solvent concentration ranges. This makes the more laborious and time-consuming isocratic experimental designs richer in information. The consequence is that transference from gradient to isocratic elution mode requires extrapolations in the solvent domain, which justifies the better results offered by the equation based on polarity parameters (Eq. (3)). This equation, which has been applied here to gradient elution, was demonstrated previously to have a particularly good performance in extrapolations under isocratic elution, for 150 compounds of very different nature (alkylbenzenes, phenols, anilines, phenones, halobenzenes, nitrobenzenes, aromatic amides, aldehydes, esters, ethers and nitriles) [9].

A practical concept proposed in this work is the informative range of solvent concentration, which is the concentration range where the accuracy of the predictions from a given gradient design is adequate. This concept can be explained reversely as the solvent concentration useful in isocratic-to-gradient predictions. The error analysis in the transference of data between elution modes, developed in this work, is proposed as a useful tool to evaluate the quality of the predictions, and eventually, detect the corresponding informative ranges. The application of these enhanced predictions to find optimal gradient separations will be considered in the future.

## Acknowledgements

This work was supported by Project BQU2001–3047 (Ministerio de Ciencia y Tecnología of Spain) and Project CTIDIB/2002/226 (Generalitat Valenciana). JRTL and GVT thank the MCYT for a Ramón *y* Cajal position, and the Generalitat Valenciana for an FPI grant, respectively.

## References

[1] L.R. Snyder, J.W. Dolan, D.C. Lommen, J. Chromatogr. 485 (1989) 65.
[2] J.W. Dolan, D.C. Lommen, L.R. Snyder, J. Chromatogr. 485 (1989) 91.
[3] R. Cela, M. Lores, Comput. Chem. 20 (1996) 175.
[4] J.R. Torres-Lapasió, M.C. García-Alvarez-Coque, J.J. Baeza-Baeza, Anal. Chim. Acta 348 (1997) 187.
[5] S. Heinisch, E. Lesellier, C. Podevin, J.L. Rocca, A. Tchapla, Chromatographia 44 (1997) 529.
[6] M. Rosés, E. Bosch, Anal. Chim. Acta 274 (1993) 147.
[7] A. Vailaya, Cs. Horváth, J. Chromatogr. A 829 (1998) 1.
[8] P. Nikitas, A. Pappa-Louisi, P. Agrafiotou, J. Chromatogr. A 946 (2002) 9.
[9] J.R. Torres-Lapasió, M.C. García-Alvarez-Coque, M. Rosés, E. Bosch, J. Chromatogr. A 955 (2002) 19.
[10] P. Jandera, J. Chromatogr. 485 (1989) 113.
[11] T. Yoshida, T. Okada, J. Chromatogr. A 841 (1999) 19.
[12] C.H. Lee, J.W. Lee, K.H. Row, J. Chromatogr. A 828 (1998) 337.
[13] J.W. Dolan, L.R. Snyder, R.G. Wolcott, P. Haber, T. Baczek, R. Kaliszan, L.C. Sander, J. Chromatogr. A 857 (1999) 41.
[14] P.J. Schoenmakers, A. Bartha, H.A.H. Billiet, J. Chromatogr. 550 (1991) 425.
[15] M.A. Quarry, R.L. Grob, L.R. Snyder, Anal. Chem. 58 (1986) 907.
[16] E. Marengo, C. Baiocchi, M.C. Gennaro, P. Bertolo, Chromatographia 27 (1989) 19.
[17] P.J. Schoenmakers, H.A.H. Billiet, R. Tijssen, L. de Galan, J. Chromatogr. 149 (1978) 519.
[18] J.R. Torres-Lapasió, M. Rosés, E. Bosch, M.C. García-Alvarez-Coque, J. Chromatogr. A 886 (2000) 31.
[19] L.R. Snyder, J.W. Dolan, Adv. Chromatogr. 38 (1998) 115.
[20] N. Lundell, J. Chromatogr. 639 (1993) 97.
[21] L.S. Madamba-Tan, J.K. Strasters, M.G. Khaledi, J. Chromatogr. A 683 (1994) 335.
[22] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in Fortran, Cambridge University Press, New York, 1994.
[23] D.L. Massart, B.G.M Vandeginste, L.M.C. Buydens, S. de Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics, Part A, Elsevier, Amsterdam, 1998.
[24] J.D. Stuart, D.D. Lisi, L.R. Snyder, J. Chromatogr. 485 (1989) 657.