Luciana Vera-Candioti[1]
María J. Culzoni[1]
Alejandro C. Olivieri[2]*
Héctor C. Goicoechea[1]

[1]Laboratorio de Desarrollo Analítico y Quimiometría, Cátedra de Química Analítica I, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Ciudad Universitaria, Santa Fe, Argentina
[2]Departamento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario, Argentina

Received May 12, 2008
Revised June 25, 2008
Accepted July 11, 2008

Research Article

Chemometric resolution of fully overlapped CE peaks: Quantitation of carbamazepine in human serum in the presence of several interferences

Drug monitoring in serum samples was performed using second-order data generated by CE-DAD, processed with a suitable chemometric strategy. Carbamazepine could be accurately quantitated in the presence of its main metabolite (carbamazepine epoxide), other therapeutic drugs (lamotrigine, phenobarbital, phenytoin, phenylephrine, ibuprofen, acetaminophen, theophylline, caffeine, acetyl salicylic acid), and additional serum endogenous components. The analytical strategy consisted of the following steps: (i) serum sample clean-up to remove matrix interferences, (ii) data pre-processing, in order to reduce the background and to correct for electrophoretic time shifts, and (iii) resolution of fully overlapped CE peaks (corresponding to carbamazepine, its metabolite, lamotrigine and unexpected serum components) by the well-known multivariate curve resolution-alternating least squares algorithm, which extracts quantitative information that can be uniquely ascribed to the analyte of interest. The analyte concentration in serum samples ranged from 2.00 to 8.00 mg/L. Mean recoveries were 102.6% ($s = 7.7$) for binary samples, and 94.8% ($s = 13.5$) for spiked serum samples, while CV (%) = 4.0 was computed for five replicate, indicative of the acceptable accuracy and precision of the proposed method.

Keywords:
Carbamazepine / Overlapped peaks / Second-order advantage
DOI 10.1002/elps.200800400

1 Introduction

CE has become the technique of choice for a wide variety of analytical applications (*e.g.* environmental, bioanalytical, pharmaceutical, food science, *etc.*) [1]. When developing a CE method, it is customary to optimize the experimental conditions, in order to guarantee complete separation of all sample components [2]. If the latter requirement is not accomplished, however, the use of chemometrics may provide a useful resource for accurate analyte quantitation [3]. Among other alternatives, this is possible if second-order data are recorded, for example, using a DAD during the electrophoretic time evolution. Interestingly, second-order data may show the intrinsic property of the second-order advantage [4], which in principle permits analyte quantitation in samples containing unexpected components, *i.e.*, compounds not included in the calibration set. This property allows one to train a predictive model with a limited number of standards, yet quantitating the analyte in the presence of potential interferents [5]. In this context, a small number of applications have been presented in which partially overlapping peaks were resolved by resorting to multivariate curve resolution-alternating least squares (MCR-ALS) [6–9], although other chemometric methodologies such as artificial neural networks [10] and wavelet transform [11] were also successfully applied.

Frequently, differences in axial diffusion can originate dissimilarities in peak shapes, leading to data without the property of trilinearity, and making necessary the use of MCR-ALS or alternatives such as PARAFAC2, a variant of PARAFAC (parallel factor analysis) allowing for distinct time profiles in each experimental sample [12]. On the other hand, an important aspect related to the analysis of

Correspondence: Professor Héctor C. Goicoechea, Cátedra de Química Analítica I, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Ciudad Universitaria, Santa Fe S3000ZAA, Argentina
E-mail: hgoico@fbcb.unl.edu.ar
Fax: +54-342-4575205

Abbreviations: AAS, acetylsalicylic acid; **CAF,** caffeine; **CBZ,** carbamazepine; **CBZ-EP,** carbamazepine epoxide; **FEN,** phenylephrine; **IBU,** ibuprofen; **LAMO,** lamotrigine; **MCR-ALS,** multivariate curve resolution-alternating least squares; **PAR,** acetaminophen; **PHB,** phenobarbital; **PHT,** phenytoin; **SVD,** singular value decomposition; **TEO,** theophylline

*Additional corresponding author: Professor Alejandro C. Olivieri, E-mail address: aolivier@fbioyf.unr.edu.ar

second-order electrophoretic-spectral data is the fact that analytes, and also potential interferents, may show different spectra but identical migration rates, for example, when sample components present similar physico-chemical properties (*e.g.* similar pK_as and molecular masses), which make the separation difficult. Interestingly, the structure of most second-order algorithms is such that they cannot deal with total overlapping in the time dimension between the analyte and the potential interferents. MCR-ALS is an algorithm that can solve this type of problems by resorting to the mathematical resource of matrix augmentation. Although we have applied the latter algorithm in a different context, *i.e.*, for the resolution of kinetic-spectral systems in which analytes and interferents present similar kinetic profiles [13], the results can be extended to the presently studied systems. Augmentation consists of constructing a grand data matrix by joining matrices for different samples, an operation which can be performed in the directions of: (i) the columns, (ii) the rows, or (iii) both columns and rows simultaneously. In electrophoretic-spectroscopic systems, it is usual to augment matrices in the time direction, *i.e.*, to place matrices on top of each other when the rows represent spectra and the column time profiles, because this alleviates the problems associated with sample-to-sample differences in this dimension. We have shown that MCR-ALS is one of the few second-order algorithms, which is able to achieve the second-order advantage, even in the presence of full overlapping between the profiles of a given analyte and potential interferents [13]. This result is of paramount importance for the present work, because the analytes and some interferences present the same electrophoretic time.

Additionally, CE instrumentation is known to present a reduced performance in injection precision, mainly owing to pressure difference, capillary length, capillary diameter, sample solution and buffer viscosities, solution surface tension, *etc.* [8]. The use of internal standards has been proposed to circumvent this problem [7, 14]. Other issues arising in CE instrumentation, which call for data preprocessing, are baseline and time shifts between runs [15]. Different strategies for baseline correction have been proposed in the literature [15]. We have employed a method proposed by Eilers, which is the multidimensional extension of the spline-based approach, a novel algorithm taking advantage of the special structure of both the data as an array and the model matrix as a tensor product [16]. In the case of time shift correction, we have implemented the evolving routine presented by Fraga *et al.*, which takes into account the second-order structure of the data [17, 18].

In the present work, we discuss for the first time the resolution of completely overlapped CE peaks, by processing CE-DAD second-order data with the MCR-ALS algorithm. We show that the analysis of carbamazepine (CBZ) can be carried out in serum samples even when other drugs occur in the samples. The method is rapid, low-cost, and shows excellent analytical performance. Several alternatives have been documented in the literature for the determination of the present analyte of interest, most of them using HPLC and also CE [19–27]. In the latter case, either a surfactant or a cyclodextrin were added to the BGE. In this report, we have used the simplest experimental conditions for the CE analysis of CBZ in the presence of a variety of other drugs and also serum components. Our purpose is to show that chemometrics can assist in this analysis, even under conditions where the electrophoretic resolution of the analyte and some sample components is not only incomplete, but the overlap in the time dimension is almost total. We feel CE researchers may benefit from the results of the present model study, learning that second-order CE-DAD data, measured from complex systems with unresolved CE peaks and uncalibrated interferents, can still be useful for successful analyte prediction. This may be helpful in avoiding complex separations systems that may require long analysis times.

2 Materials and methods

2.1 Instrumentation

All experiments were carried out on a CE system (Agilent Technologies, Waldbronn, Germany), equipped with a DAD. The instrument was operated under positive polarity. These measurements were done in an uncoated fused-silica capillary (Microsolv, USA) having 53.5 cm total length (effective length 45 cm), with an inner diameter of 75 μm. The cartridge was maintained at 25.0°C. A constant voltage of +20 kV was applied with typical current of about 90 μA. Injection was made from the positive cathode of the capillary by 50 mbar pressure for 10 s.

For each electropherogram, spectra were registered in the range 189–351 nm each 2 nm, at regular steps 0.003 s for a total time of 5 min. Therefore, matrices of size 616×82 *per* sample were generated although selected regions were subsequently employed for multivariate calibration.

With the aim of quantifying CBZ by univariate calibration, a wavelength of 214 nm was used.

2.2 Electrophoretic conditions

Careful washing of the capillary was essential in order to obtain reproducible results and to avoid interference due to the matrix and the analytes interacting with the internal wall of the capillary.

Thus, at the beginning of every working day the capillary was rinsed with: 0.1 mol/L sodium hydroxide (10 min), Milli-Q water (10 min), and 0.025 mol/L sodium borate buffer pH 9.20 (10 min), which is the BGE. Between runs the capillary was flushed successively with 0.1 mol/L sodium hydroxide, Milli-Q water, and BGE for 3 min each. At the end of the day the capillary was washed with 0.1 mol/L sodium hydroxide (5 min), Milli-Q water (5 min), and then with air-dried for 3 min.

All the solutions were degassed in an ultrasonic bath and filtered through 0.45 μm nylon membrane (Sartorius, Germany) before use.

2.3 Software

In all cases, electropherograms were recorded in random order with respect to analyte concentrations, and those corresponding to the calibration set were recorded in different days with respect to the test set. Data were saved in ASCII format, and transferred to a PC Sempron AMD microcomputer for subsequent manipulation by chemometric programs. MCR-ALS was implemented using the graphical interface provided by R. Tauler in his web page (http://www.ub.edu/mcr/welcome.html) [28]. Baseline correction was done with a MATLAB code written by Eilers [16]. Time shift correction was applied with a MATLAB code written according to Ref. [17]. All the applications were run as MATLAB 7.0 routines (MATLAB 7.0, The Mathworks, Natick, Massachusetts, 2003).

2.4 Reagents

CBZ (5H-dibenz[b,f]azepine-5-carboxamide), its main metabolite carbamazepine epoxide (1a,10b-dihydro-6H-dibenzo-[b,f]oxireno[d]azepine-6-carboxamide, CBZ-EP), phenobarbital (5-ethyl-5-[1-methylbutyl]-2,4,6-trioxohexahydropyrimidine], PHB), phenytoin (5,5-diphenylhydantoin, PHT), phenylephrine (3-(1-hydroxy-2-methylamino-ethyl)-phenol, FEN), ibuprofen (α-methyl-4-[isobutyl] phenylacetic acid, IBU), acetaminophen (4-acetamidophenol, PAR), theophylline (1,3-dimethylxanthine, TEO), caffeine (1,3,7-trimethylxanthine, CAF), and acetylsalicylic acid (2-acetoxybenzoic acid, AAS) were provided by Sigma-Aldrich. Lamotrigine (6-(2,3-dichlorophenyl)-1,2,4-triazine-3,5-diamine, LAMO) was provided by Química Alkano S.A. (México). Sodium borate, sodium phosphate, sodium hydroxide, dichloromethane, and methanol were analytical grade. Milli-Q quality water was used in all the CE experiments.

Stock solutions were prepared in the following solvents: CBZ, CBZ-EP, and LAMO in methanol; FEN, PHB, PHT, PAR, AAS, CAF, and TEO in water and IBU in ethanol, at concentrations of 2 g/L, and stored in the darkness at −4°C. Standard working solutions of all analytes were prepared everyday by dilution of the stock solutions in BGE.

2.5 Extraction procedure

A given serum sample (500 μL) was transferred to a round bottom polypropylene tube, and the analytes were extracted with three successive aliquots of 1.00 mL of dichloromethane. The tube was capped, then shaken by vortex and afterwards centrifuged at 4000 rpm during 5 min. The lower organic layer was transferred to a clean tube and evaporated to dryness at 45°C in a hot plate. The whole operation takes *ca.* 20 min, which is significantly less than the time required for other extraction procedures reported in the literature [19–27].

Finally, the residue was redissolved in 500 μL of BGE and filtered through a syringe filter of 0.45 μm pore size prior to injection into the CE instrument.

2.6 Calibration and validation sets

For calibration, a set of 21 standard samples was prepared by dilution of known amounts of CBZ stock solution in BGE. Seven concentration levels in the range from 1.00 to 15.00 mg/L were prepared in triplicate and were analyzed by the CE system. This range included the average therapeutic serum concentration of CBZ reported in the literature (from 4.00 to 12.00 mg/L) [29].

For computing figures of merit under univariate calibration, peak areas of CBZ were plotted against the nominal concentrations (expressed in mg/L) and the line was fitted by a least-squares method. The values of LOQ and LOD were calculated according to IUPAC recommendations [30].

Several test samples containing mixtures of CBZ and CBZ-EP at concentration levels different than those for calibration were prepared by dilution of standard working solutions in BGE (B1–B5, see Table 1).

In addition, 15-spiked serum test samples were prepared by adding known amounts of CBZ and CBZ-EP

Table 1. Composition of the test samples

Sample[a)]	CBZ (mg/L)	CBZ-EP (mg/L)	LAMO (mg/L)
B1	2.50	2.50	–
B2	2.98	5.99	–
B3	3.97	7.99	–
B4	6.99	2.00	–
B5	7.50	7.38	–
SA1	3.02	5.99	–
SA2	3.02	7.03	–
SA3	3.97	2.00	–
SA4	3.97	7.99	–
SA5	6.03	5.99	–
SB1	2.00	8.00	–
SB2	3.00	5.00	–
SB3	4.00	6.00	–
SB4	6.00	4.00	–
SB5	8.00	2.00	–
SC1	4.96	4.99	5.66
SC2	4.96	4.99	7.55
SC3	4.96	4.99	9.44
SC4	4.96	4.99	11.33
SC5	4.96	4.99	13.21

a) The first capital letter corresponds to the type of matrix in which each sample was prepared: B, BGE; S, serum. The second capital letter identifies other components present in each sample: A, CBZ-EP; B, CBZ-EP and serum unexpected components; C, CBZ-EP and LAMO.

standard solutions to blank sera, obtained from healthy patients who are not under anti-epileptic treatment (SA1–SA5, SB1–SB5, and SC1–SC5), five of them with the addition of known concentrations of LAMO (SC1–SC5, see Table 1). The resulting mixtures were subjected to the extraction procedure. Several spiked serum samples containing the same concentration of CBZ were used to evaluate precision.

On the other hand, accuracy was evaluated by means of recovery assays. Known amounts of CBZ standard solutions were added to 500 μL of a healthy patient serum sample in order to reach concentrations of 2.00, 4.00, 6.00, and 8.00 mg/L. These mixtures were subjected to the extraction procedure and then analyzed in the CE system.

An additional sample was prepared for use in an interferent study. It consisted of a serum sample, which was spiked with the following drugs: (i) CBZ and CBZ-EP (5.00 mg/L), (ii) FEN (20.00 mg/L), and (iii) LAMO, PAR, IBU, PHT, PHB, TEO, CAF, and AAS (10.00 mg/L).

3 Theory

3.1 Baseline correction

The implementation of a properly pre-processing step should be considered of an extreme importance, in order to guarantee the quality of instrumental signals. This fact determines the quality of the final results of the electrophoretic analysis. As with any other instrumental signal, electropherograms are composed of three major components: signal, noise, and background, which differ in their frequency. Consequently, eliminating noise and background components may result in a signal enhancement [15].

Moreover, elimination of the electropherogram baseline is a critical step for reducing the number and complexity of the unexpected components. With this aim, we applied the methodology proposed by Eilers, *i.e.* the asymmetric least-squares method [31], which was recently adapted to multi-dimensional data [16]. This method consists in the estimation of the background matrix from the matrix data, using spline basis matrices along their columns and rows [32].

3.2 Peak shift correction

The methodology proposed by Prazen and coworkers [18] was employed, which takes into account the second-order structure of the data. The algorithm is based on the singular value decomposition (SVD) of an **N**/**M** matrix joining the data matrices **N** and **M**, where **N** is taken as reference and **M** is to be corrected in relation to **N**:

$$(\mathbf{U}, \mathbf{S}, \mathbf{V}) = \mathrm{SVD}(\mathbf{N}/\mathbf{M}) \tag{1}$$

In Eq. (1), **S** is a diagonal matrix containing the singular values, **U** and **V** are matrices containing the left and right singular vectors of **N**/**M**, respectively. The correction is done by computing the residual variance (RES) while the matrix **M** is moved in relation to **N**, using a pre-established number of points that can be estimated by the inspection of the electropherograms of **N** and **M**.

$$\mathrm{RES} = 100 \frac{\sum_{a=A+1}^{\min(J,K)} s_a^2}{\sum_{a=1}^{\min(J,K)} s_a^2} \frac{JK}{(J-A)(K-A)} \tag{2}$$

where A is the number of significant singular values, ideally equal to the number of species presents in **N**/**M**, and J and K are the number of digitized warelengths and the number of migration times, respectively. When the matrices **N** and **M** are aligned, the RES values should reach a minimum.

3.3 MCR-ALS

As was previously mentioned, this method is capable of dealing with data sets deviating from trilinearity. Instead of forming a three-dimensional data array, the latter is unfolded along the mode that is suspected of breaking the trilinear structure, *i.e.* if a matrix-to-matrix variation of profiles occurs along the column direction, a column-wise augmented matrix is created. The bilinear decomposition of the augmented matrix **D** is performed according to the expression:

$$\mathbf{D} = \mathbf{C} \times \mathbf{S}^{\mathrm{T}} + \mathbf{E} \tag{3}$$

where the rows of **D** contain the absorption spectra measured as a function of time, the columns of **C** contain the time profiles of the compounds involved in the process, the columns of **S** their related spectra, and **E** is a matrix of residuals not fitted by the model. Appropriate dimensions of **D**, **C**, **S**, and **E** are thus $(1+I)K \times J$, $(1+I)K$, $J \times K$ and $(1+I)K \times J$, respectively (I = number of training samples). Decomposition of **D** is achieved by iterative least-squares minimization of $||\mathbf{E}||$ under suitable constraining conditions, *i.e.* nonnegativity in spectral profiles, unimodality and nonnegativity in concentration profiles.

The pure spectra of the compounds should be the same in all experiments, but the profiles in the different **C** submatrices need not share a common shape. This is the reason why electrophoretic runs, which are affected by the previously mentioned instrumental pitfall and also kinetic experiments performed in different conditions (*e.g.* pH, temperature, *etc.*) and, hence, showing different kinetic profiles, can be analyzed together as long as the spectra of the compounds involved in the process remain invariant (absorption spectra usually show a limited temperature dependence).

In the present context, it is necessary to point out that MCR-ALS requires initialization with system parameters as close as possible to the final results. In the column-wise augmentation mode, the species spectra are required, as obtained from either pure analyte standards or from the analysis of the purest spectra based on the so-called SIMPLISMA (simple interactive self-modeling mixture analysis) methodology [33], an MCR algorithm, which

extracts pure component spectra from a series of spectra of mixtures of varying composition. In the case of truly unknown samples with possible interferents, we have found that successful application of MCR-ALS to our case required the combination of both of the above procedures, *i.e.*, known spectra for pure analytes and purest spectral analysis for the interferent.

Finally, successful MCR is also aided by the inclusion of information about the correspondence among species in each matrix (*i.e.*, information as to whether a given component exists or not in a given sample) [36].

4 Results and discussion

4.1 Electrophoretic analysis

Several authors have reported the successful separation of CBZ and CBZ-EP in serum samples by HPLC [19–22] and also by MECK [23–27]. They have made great efforts to achieve not only the mutual separation of both analytes, but also their distinction from other anti-epileptic drugs. Most of these works required rather long analysis times (*ca.* 15–20 min).

In the present work, in order to find the best conditions leading to a correct peak shape for CBZ in a minimum analysis time, a screening phase was carried out. The capabilities of six BGE were evaluated: sodium borate pH 9.20 and sodium phosphate pH 4.2, at three concentration levels: 0.01, 0.02, and 0.03 mol/L. The evaluation consisted in analyzing a standard working solution of CBZ (5.00 mg/L) diluted in water, methanol, and BGE.

According to the results from the screening phase, a sodium borate buffer 0.025 mol/L (pH 9.20) was selected due to the provision of: (i) satisfactory peak shape for CBZ (without appreciable deformations), (ii) minimization of the analysis time, reaching migration times of 3.50 min, and (iii) minimization of the baseline noise. Eighteen electropherograms were obtained combining the different variables being analyzed, and their visual inspection allowed us to select the most convenient experimental conditions. Standard working solutions and serum sample extracts were diluted in BGE because it provides larger peak areas.

4.1.1 Univariate calibration and figures of merit

In order to verify the method linearity within a concentration range from 1.00 to 15.00 mg/L of CBZ, a least-squares fitting was performed with the obtained data (peak areas *versus* concentrations), and the results are presented in Table 2. The performed ANOVA test of lack of fit allows one to conclude that linearity is fulfilled within the studied range.

The LOD (0.9 mg/L) and the LOQ (2.8 mg/L) obtained in this work are low enough to determine and quantitate therapeutic concentrations of CBZ in patients treated with doses from 200 to 1200 mg of CBZ, which are known to give rise to CBZ serum levels in the 4.00–12.00 mg/L range. The LOD was obtained using the IUPAC's recommended procedure based on false positive and false negative probabilities (both 0.05), as described in Ref. [34].

Table 2. Univariate figures of merit for the linearity assay

Computed parameters	
r-Squared	0.9946
Intercept[a)]	0 (2)
Slope[a)]	13.1 (0.2)
p-Value (lack of fit)[b)]	0.522
LOD (mg/L)	0.9
LOQ (mg/L)	2.8

a) Values between parentheses are standard deviations.
b) Considered significant when $p<0.05$.

Table 3. Extraction yield for spiked serum samples

Sample	CBZ spiked to blank serum (mg/L)	Recovery (%)
1	2.00	98.7
2	4.00	103.2
3	6.00	99.2
4	8.00	101.5
Mean recovery (%)	–	100.6

On the other hand, the absolute recovery was evaluated on blank serum samples spiked with four different concentrations of CBZ (see Section 2). The analyte concentrations in spiked samples were obtained by interpolation on the corresponding calibration curve, and then compared with the theoretical amount added. The average recovery value of 100.6% given in Table 3 is indicative of the high accuracy obtained in the four concentration levels studied considering the requirements for bioanalytical assays [35].

4.2 Baseline and time shift correction

With the object of simplifying the analysis, both background elimination and time shift correction were performed on the second-order electrophoretic-spectral data. Both data preprocessing procedures were implemented by using algorithms that were previously adapted by their authors to multidimensional data [16, 18].

For obtaining the **F** background correction matrix, ten spline basis function were used, with a single regularization parameter, whose value was 1. Figure 1A shows the landscape corresponding to a serum sample (SC1 in Table 1) that was spiked with CBZ, CBZ-EP, and LAMO. As can be seen in this figure, there is a single peak in the time domain, indicating the complete overlapping of the three substances. Figure 1B shows the corresponding background matrix (**F**).

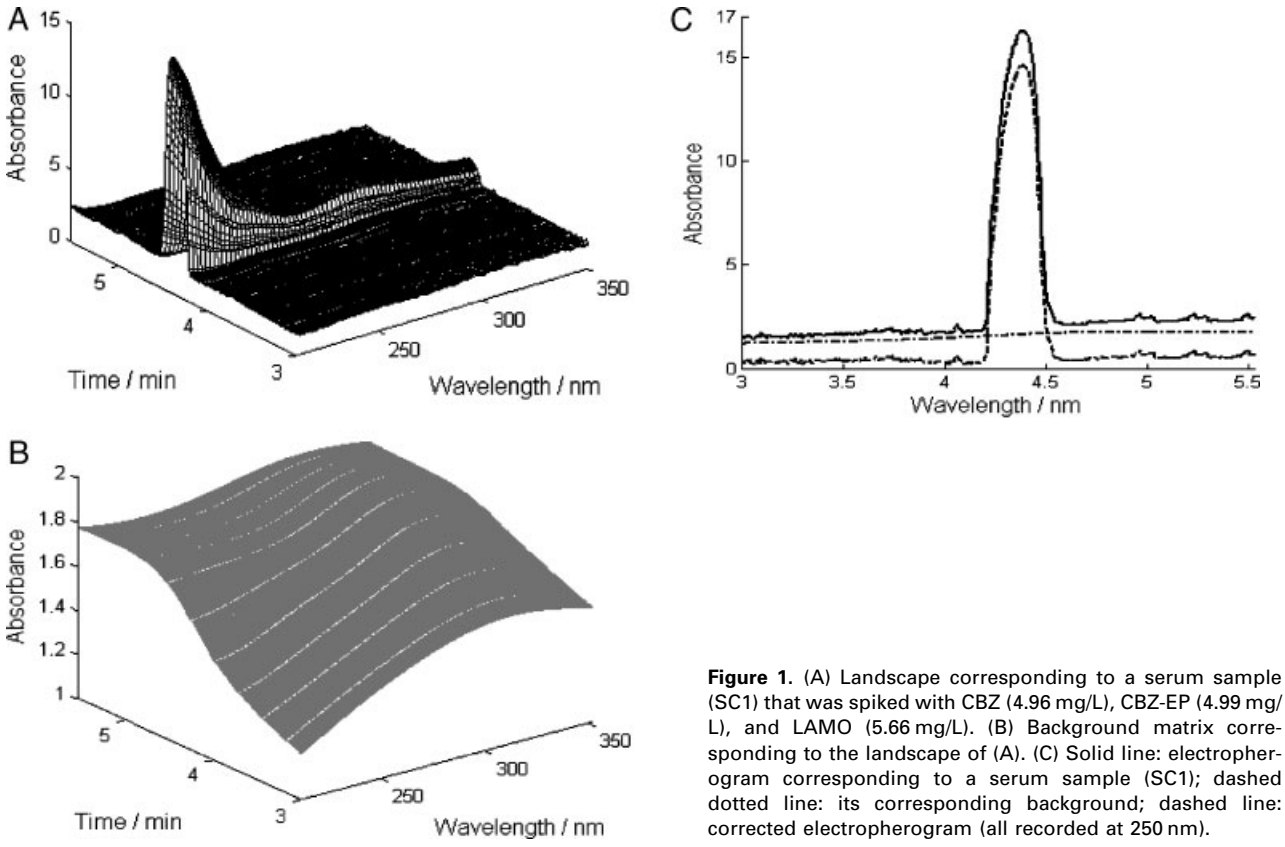

Figure 1. (A) Landscape corresponding to a serum sample (SC1) that was spiked with CBZ (4.96 mg/L), CBZ-EP (4.99 mg/L), and LAMO (5.66 mg/L). (B) Background matrix corresponding to the landscape of (A). (C) Solid line: electropherogram corresponding to a serum sample (SC1); dashed dotted line: its corresponding background; dashed line: corrected electropherogram (all recorded at 250 nm).

A different scale was used in order to show the variation of the baseline at different wavelengths. Notably, the electrophoretic baseline increases with wavelength until a maximum at *ca*. 300 nm is reached. This fact justifies using a second-order background correction algorithm. Subtraction of this matrix allowed us to obtain a corrected data matrix, which was subsequently submitted to time shift correction. Finally, Fig. 1C shows the original electropherogram plotted at 250 nm, and the corresponding baseline and corrected electropherogram.

The need of time shift correction can be appreciated in Fig. 2A, where the electropherograms recorded at 250 nm for five CBZ standards and one serum spiked sample (the same which was selected for illustrating the background correction) are plotted. The time shift is evident, and although MCR is able to deal with non-synchronized data, subsequent analysis showed that correction improves the results. As previously mentioned, correction of a matrix **M** (solid lines in Fig. 2A), taking a matrix **N** as a reference (plotted as short dotted lines), consisted in performing SVD to the joint **N**/**M** data matrix. The number of necessary components ranged between 2 and 3, depending on the matrix being analyzed. Also, the window was varied in order to optimize the application. As an example, Fig. 2B shows the RES(%) variation (see Eq. 2) when the studied matrix is shifted with respect to the reference matrix. The most remote peak needed a window correction of 39 time points to reach a minimum. Finally, Fig. 2C shows electropherograms recorded at 250 nm for all the six aligned peaks, with respect to the reference. It is interesting to note that although peaks look approximately aligned, the loss in trilinearity is evident, because the shapes of the individual bands are all different, making it necessary the application of the MCR algorithm.

4.3 MCR-ALS application to mixtures in BGE and to spiked serum samples

Figure 3A shows a complete electropherogram of a BGE sample having not only CBZ and its main metabolite CBZ-EP, but also other anti-epileptic drugs such as PHT, PHB, and LAMO, which are likely to be co-administered with CBZ, as well as a variety of other commonly used pharmaceuticals: FEN, PAR, TEO, IBU, CAF, and AAS. As can be seen, the electrophoretic separation is not complete, yielding the overlapping of CBZ, CBZ-EP, CAF, and LAMO (peak 2) and both TEO and IBU (peak 5). Furthermore, the extraction procedure applied to the sample recovers three of the four analytes included in peak 2 (CBZ, CBZ-EP, and LAMO), making impossible the determination of CBZ by univariate calibration (see Fig. 3B). Although several other analytes were also extracted, they were not included in the experiments since they are not potential interferences, due to a complete separation from CBZ.

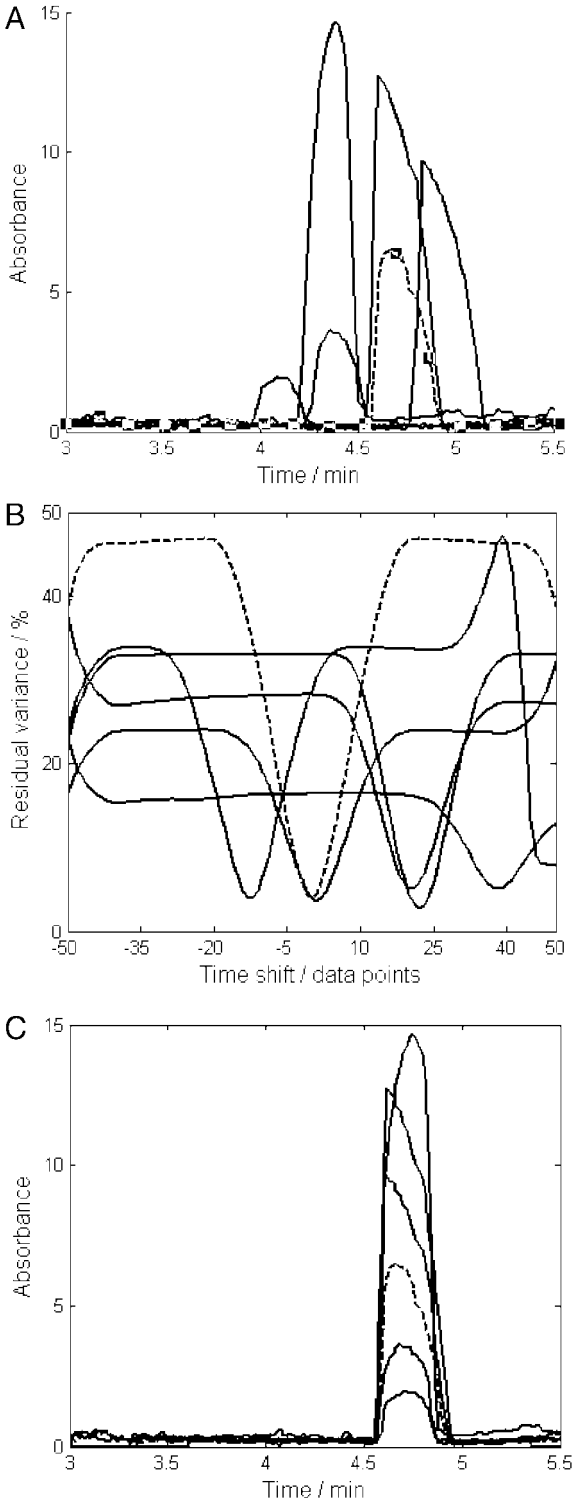

Figure 2. (A) Electropherograms recorded at 250 nm for five CBZ standards and one spiked serum sample (SC1). (B) RES (%) variation (see Eq. 2) when the studied matrix is shifted with respect to the reference matrix. (C) Shift corrected electropherograms recorded at 250 nm.

4.3.1 Exploratory rank analysis of the data sets

In MCR-ALS, the performance of the resolution strongly depends on the knowledge of the global and local properties of the data set, particularly on those related to the mathematical and chemical rank [36]. In practice, a data matrix is considered full rank if the estimated rank is equal to the number of absorbing chemical species, assuming that this number is known [36].

The first step in the application of MCR is the determination of the correct number of compounds in the studied matrix, using, for instance, the SVD technique [37]. In the present case, when single data matrices containing three chemical species having almost identical electrophoretic times (CBZ, CBZ-EP, and LAMO) were analyzed, SVD was not capable of detecting the right number of contributing components, which explain the variance of the system, due to their identical time profiles, leading to rank-deficient matrices. Although some selectivity exists in the spectral domain, the almost identical time profiles lead to the existence of a single significant principal component for this system.

To overcome the above-discussed rank-deficiency problem, matrix augmentation can be applied. This is a simple and smart way of breaking rank-deficiencies and thus facilitating resolution [38]. The strategy employed in this work was to augment the rank-deficient matrix under scrutiny with matrices for pure analyte standards, and then estimate the rank of the augmented matrix by SVD. A typical case is presented in Table 4 for sample SC3. The application of SVD to the individual serum sample matrix SC3 containing CBZ, CBZ-EP, and LAMO (Table 1) does only detect a single significant singular value instead of 3, which is the actual number of chemical components in the sample (first entry in Table 4). Only when the matrix is augmented by adding information on the analyte, CBZ-EP, and LAMO standards, application of SVD to the augmented matrix was capable of correctly estimating the rank, *i.e.*, 3 (last entry in Table 4). Furthermore, this approach helped us to detect the presence of unexpected components in some of the serum samples (*i.e.*, the SB set, see Table 1). Since the studied serum samples were all different, some of them carry components that do show during MCR analysis, while others did not produce significant interference.

4.3.2 Multivariate calibration results

Before the resolution by MCR-ALS, baseline and time shift correction was applied to all data matrices, except to the SB set, which contains CBZ and CBZ-EP. As can be seen in Fig. 4, the electropherogram of an SB sample contains a broad peak corresponding to serum components eluting together with both CBZ and CBZ-EP. This fact allowed us to only apply background correction, making it impossible the alignment of peak 2. To perform the MCR quantitation taking advantage of its capability to handle non-trilinear data, appropriate time regions were selected for SB samples,

according to the position of the peak of interest in each of the electropherograms corresponding to each of the matrices involved in the augmented data matrix.

Each of the test samples was analyzed as follows. The spectral-time matrix for a given test sample was augmented with the 21 matrices recorded for the calibration samples (only containing the analyte CBZ). An additional matrix, recorded for a standard of CBZ-EP was joined with the augmented matrix in all cases. Finally, in the case of all SC serum samples, a matrix recorded for a standard of LAMO was added. The CBZ-EP and LAMO matrices were added in order to provide sufficient information to MCR for the successful resolution of the sample components. Table 5 presents the results obtained for CBZ when the test samples were analyzed. It may be noticed that SVD analysis detected two components for sets B and SA, *i.e.*, one for CBZ and the other one for CBZ-EP, whereas an additional component was required for sets SB and SC, corresponding to an unexpected serum component and LAMO, respectively. For the set of samples B, SA, and SC, which contain CBZ and CBZ-EP (and also LAMO in the latter case), the initial estimates were built by resorting to the pure standard spectra for each analyte. Those for the SB set also included an unexpected component profile obtained by the application of SIMPLISMA [38] to an SB sample. The constraints imposed during the ALS procedure were non-negativity in both concentration profiles and spectra, unimodality in concentration profiles, and species correspondence (see Section 3).

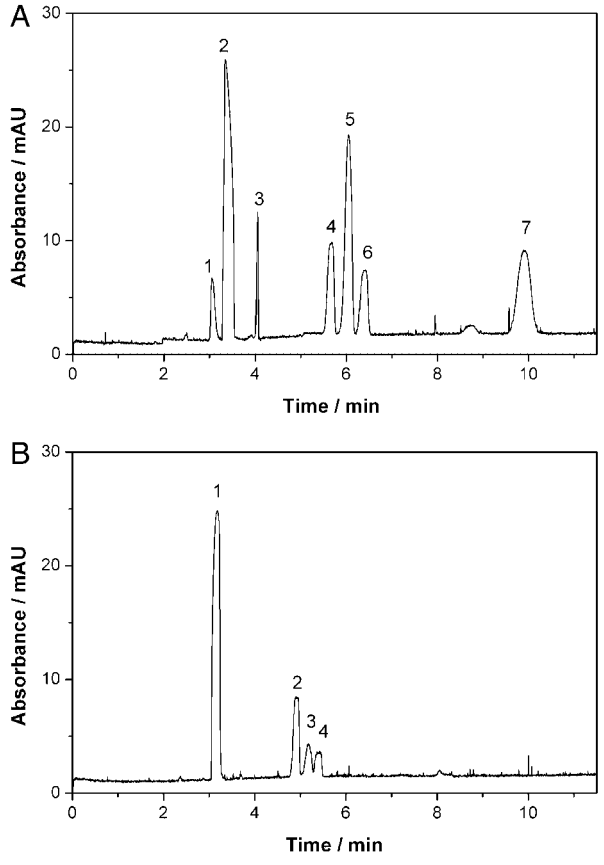

Figure 3. (A) Electropherogram (recorded at 250 nm) of a spiked BGE sample in which peak is assigned to: (1) FEN (20.00 mg/L); (2) CBZ, CBZ-EP (5.00 mg/L), CAF, and LAMO (10.00 mg/L); (3) PAR (10.00 mg/L); (4) PHT (10.00 mg/L); (5) IBU and TEO (10.00 mg/L); (6) PHB (10.00 mg/L); (7) AAS (10.00 mg/L). (B) Electropherogram (recorded at 250 nm) of the same spiked BGE sample after the extraction procedure: (1) CBZ, CBZ-EP, and LAMO; (2) PHT; (3) IBU and TEO; (4) PHB.

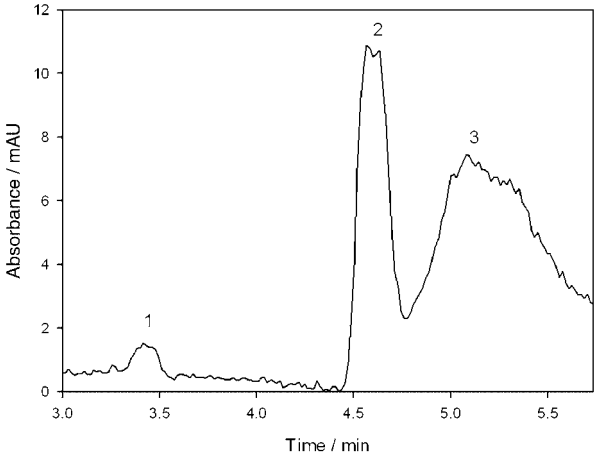

Figure 4. Electropherogram (recorded at 250 nm) corresponding to an SB serum sample containing CBZ and CBZ-EP overlapped in peak 2, and unexpected components in peaks 1 and 3.

Table 4. Estimated rank computed from singular value decomposition of the individual and augmented data sets

Augmented matrix[a)]	Known chemical rank	SVD estimated rank in the augmented matrix	Condition
[SC3]	3	1	Rank-deficient
[SC3;C]	3	2	Rank-deficient
[SC3;E;C]	3	2	Rank-deficient
[SC3;L;C]	3	2	Rank-deficient
[SC3;E;L;C]	3	3	Full-rank

a) The nomenclature $[X_1,X_2,\ldots,X_N]$ implies that an augmented **D** matrix (see Eq. 3) is created by placing matrices X_1, $X_2,\ldots,X_N$ on top of each other. It is assumed that the column direction is the electrophoretic time, while the row direction is the spectral wavelength. SC3 is a serum sample matrix containing CBZ, CBZ-EP, and LAMO (see Table 1), while C, E, and L are pure standard matrices of CBZ, CBZ-EP, and LAMO, respectively.

Figures 5A–D and 6A–C show the time and spectral profiles corresponding to the three components found by MCR-ALS analysis for an SC sample. As can be seen, overlapping of profiles in both dimensions is severe; however, MCR-ALS is able to correctly decompose the data into the relevant contributions. This allows for isolation of the signal, which can be ascribed to the analyte in each studied test sample, and therefore for accurate analyte quantitation. In addition, Fig. 6A–C shows the comparison between the spectral profiles retrieved by MCR-ALS analysis and those experimentally recorded for standards of CBZ, CBZ-EP, and LAMO. As can be seen, reasonably good agreement is found between the spectra corresponding to the analyte and interferences (correlation coefficients equal to 0.9857, 0.9843, and 0.9971, respectively). There exists high similarity between the spectra for CBZ-EP and LAMO, a fact that could explain the problems found when estimating the matrix ranks for the systems containing these three substances.

Table 5. Test concentrations and MCR-ALS prediction results

Sample[a)]	CBZ concentration (mg/L)		Recovery (%)
	Nominal	Predicted	
B1	2.50	2.42	96.8
B2	2.98	2.81	94.3
B3	3.97	4.47	112.6
B4	6.99	7.07	101.1
B5	7.50	8.13	108.4
SA1	3.02	2.30	76.2
SA2	3.02	2.91	96.4
SA3	3.97	3.58	90.2
SA4	3.97	4.23	106.5
SA5	6.03	3.86	64.0
SB1	2.00	1.92	95.8
SB2	3.00	2.44	81.4
SB3	4.00	4.42	110.4
SB4	6.00	5.52	92.0
SB5	8.00	9.15	114.4
SC1	4.96	4.71	95.0
SC2	4.96	4.98	100.4
SC3	4.96	4.64	93.5
SC4	4.96	4.53	91.3
SC5	4.96	4.53	91.3

a) The first capital letter corresponds to the type of matrix in which each sample was prepared: B, BGE; S, serum. The second capital letter identifies other components present in each sample: A, CBZ-EP; B, CBZ-EP and serum unexpected component; C, CBZ-EP and LAMO.

The prediction results are very good and are included in the acceptable range suggested by the ICH [35]. Furthermore, those given for the SC set, which is composed of five samples having the same CBZ concentration, show not only excellent recoveries but also a very good reproducibility owing to the low coefficient of variation achieved, *i.e.*, CV(%) = 4.0%.

5 Concluding remarks

Second-order data, generated by recording spectra during the electrophoretic time evolution of fully overlapped CE peaks, can be successfully handled to assess the content of CBZ in human serum, even in the presence of several interferences. This can be done using a strategy that includes data pretreatment for time shift and background correction, and data modeling with the MCR-ALS algorithm. Most of the previous literature reports on the subject required rather long analysis times (*ca.* 15–20 min), while in the present report accurate results can be achieved in only 5 min.

Universidad Nacional del Litoral (Project CAI+D No. 12/B604), Universidad Nacional de Rosario, CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas, Project No. PIP 5303), and ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica, Project Nos. PICT04-25825 and PICTO05-35124) are gratefully acknowledged for financial

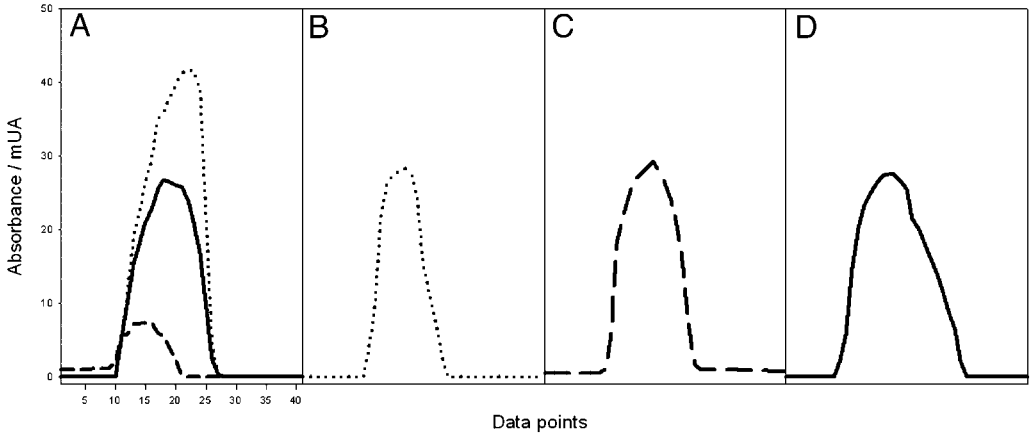

Figure 5. Time profiles retrieved by MCR-ALS when processing an SC sample, corresponding to: (A) matrix SC, the three found components; (B) matrix CBZ-EP standard (5.00 mg/L); (C) matrix LAMO standard (5.00 mg/L); (D) matrix CBZ standard (5.00 mg/L).

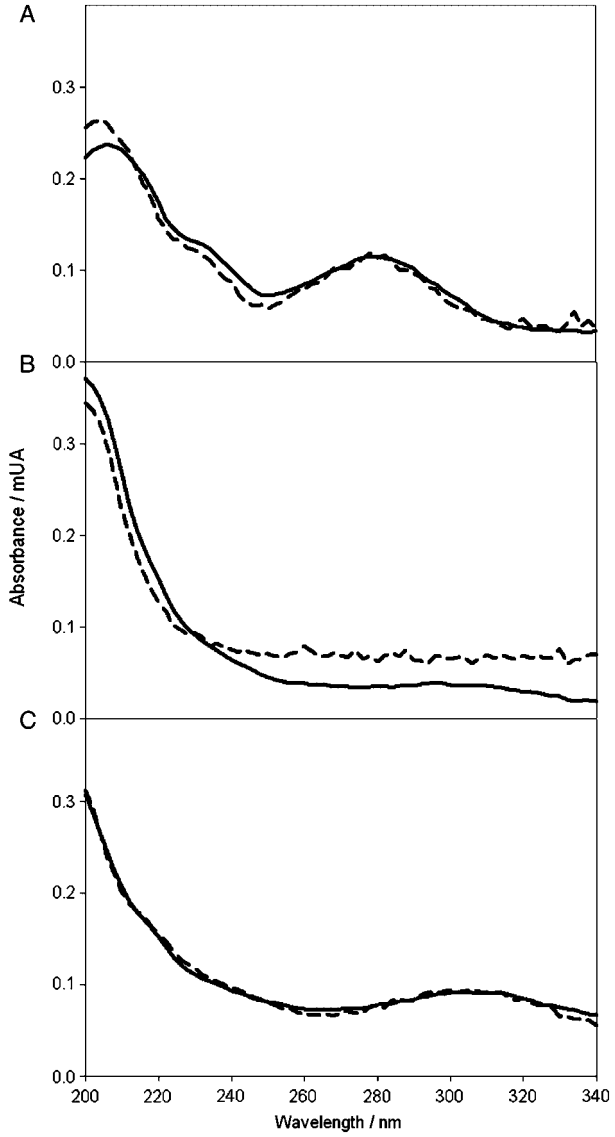

Figure 6. Comparison of spectral profiles retrieved by MCR-ALS when processing the SC sample analyzed in Fig. 5. In all the three cases, pure standard spectra are presented as dashed lines, whereas the extracted profiles are shown as solid lines: (A) CBZ; (B) CBZ-EP; (C) LAMO.

support. L. V.-C. and M. J. C. thank CONICET for their fellowships. The authors are very grateful to P. H. C. Eilers and J. W. Braga for the MATLAB codes.

The authors have declared no conflict of interest.

6 References

[1] Kraly, J., Abul Fazal, M. D., Schoenherr, R. M., Bonn, R. *et al.*, *Anal. Chem.* 2006, *78*, 4097–4110.

[2] Hanrahan, G., Montes, R., Gomez, F. A., *Anal. Bioanal. Chem.* 2008, *390*, 169–179.

[3] Sentellas, S., Saurina, J., *Electrophoresis* 2003, *26*, 1395–1402.

[4] Booksh, K. S., Kowalski, B. R., *Anal. Chem.* 1994, *66*, 782A-791A.

[5] Escandar, G. M., Faber, N. M., Goicoechea, H. C., Muñoz de la Peña, A. *et al.*, *Trends Anal. Chem. (TRAC)* 2007, *26*, 752-765.

[6] Sentellas, S., Saurina, J., Hernández-Cassou, S., Galceran, M. T., Puignou, L., *Anal. Chim. Acta* 2001, *431*, 49–58.

[7] Li, H., Zhang, F., Havel, J., *Electrophoresis* 2003, *24*, 3107–3115.

[8] Zhang, F., Li, H., *Chemom. Intell. Lab. Syst.* 2006, *82*, 184–192.

[9] Li, H., Hou, J., Wang, K., Zhang, F., *Talanta* 2006, *70*, 336–343.

[10] Zhang, Y., Li, H., Hou, A., Havel, J., *Chemom. Intell. Lab. Syst.* 2006, *82*, 165–175.

[11] Jiao, L., Gao, S., Zhang, F., Li, H., *Talanta* 2008, *75*, 1061–1067.

[12] Bro, R., *Doctoral Thesis*, University of Amsterdam, The Netherlands, 1998.

[13] Culzoni, M. J., Goicoechea, H. C., Ibañez, G. A., Lozano, V. A. *et al.*, *Anal. Chim. Acta* 2008, *614*, 46–57.

[14] Dehouck, P., Jaggavarapu, P. K. R., Desmedt, A., Van Schpdael, A., Hoogmartens, J., *Electrophoresis* 2004, *25*, 3313–3321.

[15] Daszykowski, M., Walczak, B., *Trends Anal. Chem. (TRAC)* 2006, *25*, 1081–1096.

[16] Eilers, P. H. C., Currie, I. D., Durbán, M., *Comp. Stat. Data Anal.* 2006, *50*, 61–76.

[17] Prazen, B. J., Synovec, R. E., Kowalski, R. B., *Anal. Chem.* 1998, *70*, 218–225.

[18] Fraga, C. G., Prazen, B. J., Synovec, R. E., *Anal. Chem.* 2001, *73*, 5833–5840.

[19] Franceschi, L., Furlanut, M., *Pharm. Res.* 2005, *51*, 297–302.

[20] Batí, M. M., Hanson, G. D., Schultlz, L., *J. Pharm. Biomed. Anal.* 1998, *16*, 1233-1240.

[21] Mandrioli, R., Fiorenzo, A., Casamenti, G., Sabbioni, C., Raggi, M. A., *J. Chromatogr. B* 2001, *762*, 109–116.

[22] Bugamelli, F., Sabbioni, C., Mandrioli, R., Kenndler, E. *et al.*, *Anal. Chim. Acta* 2002, *472*, 1–10.

[23] Härtter, S., Jensen, B., Hiemke, C., Leal, M. *et al.*, *J. Chromatogr. B* 1998, *712*, 253–258.

[24] Kataoka, Y., Makino, K., Oishi, R., *Electrophoresis* 1998, *19*, 2856–2860.

[25] Raggi, M. A., Pucci, V., Mauricio, A., Muzikar, J., Kenndler, E., *J. Chromatogr. B* 2002, *770*, 217–225.

[26] Lancas, F. M., Sozza, M. A., Queiroz, M. E. C., *J. Anal. Toxicol.* 2003, *27*, 304–308.

[27] Kuldvee, R., Thormann, W., *Electrophoresis* 2001, *22*, 1345–1355.

[28] Jaumot, J., Gargallo, R., de Juan, A., Tauler, R., *Chemom. Intell. Lab. Syst.* 2005, *76*, 101–110.

[29] Smith, C. M., Reynard, A. M., *Farmacologia*, Médica Panamericana, Argentina 1993.

[30] Danzer, K., Currie, L. A., *Pure Appl. Chem.* 1998, *70*, 993–1014.

[31] Eilers, P. H. C., *Anal. Chem.* 2004, *76*, 404–411.

[32] Kaczmarek, K., Walczak, B., de Jong, S., Vandginste, B. G. M., *Acta Chromatogr.* 2005, *15*, 82–95.

[33] Windig, W., Guilment, J., *Anal. Chem.* 1991, *63*, 1425–1342.

[34] Currie, L. A., *Pure Appl. Chem.* 1995, *67*, 1699–1723.

[35] ICH, Harmonised Tripartite Guideline, Q2A and Q2B, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceutical for Human Use, 1994.

[36] Ruckebusch, C., De Juan, A., Duponchel, L., Huvenne, J. P., *Chemom. Intell. Lab. Syst.* 2006, *80*, 209–214.

[37] Muñoz, G., De Juan, A., *Anal. Chim. Acta* 2007, *595*, 198–208.

[38] De Juan, A., Casassas, E., Tauler, R., in: Meyers, R. A. (Ed.), *Encyclopedia of Analytical Chemistry*, Wiley, Chichester 2000, pp. 9800–9837.

