Anal Bioanal Chem (2002) 374 : 451–459
DOI 10.1007/s00216-002-1466-9

ORIGINAL PAPER

Juan A. Arancibia · Alejandro C. Olivieri
Graciela M. Escandar

First- and second-order multivariate calibration applied to biological samples: determination of anti-inflammatories in serum and urine

Received: 6 March 2002 / Revised: 19 June 2002 / Accepted: 25 June 2002 / Published online: 11 September 2002

Abstract First- and second-order multivariate calibration of fluorescence data have been compared as regards the determination of anti-inflammatories and metabolites in the biological fluids serum and urine. The simultaneous resolution of naproxen–salicylic acid mixtures in serum and naproxen–salicylic acid–salicyluric acid mixtures in urine was accomplished and employed for a discussion of the relative advantages of the applied chemometric tools. The analysis of second-order fluorescence excitation-emission matrices was performed using iteratively reweighted generalized rank annihilation method (IRGRAM), parallel factor analysis (PARAFAC), and self-weighted alternating trilinear decomposition (SWATLD). The results were compared with first-order fluorescence emission data analyzed with partial least-squares regression (PLS). In all cases, the performance of the methods was improved through the formation of inclusion complexes of the analytes with β-cyclodextrin. The concentration ranges in which the analytes could be determined were as follows: naproxen, 0–250 ng mL^{-1} in serum and 0–200 ng mL^{-1} in urine; salicylic acid, 0–500 ng mL^{-1} in serum and 0–300 ng mL^{-1} in urine, and salicyluric acid, 0–300 ng mL^{-1} in urine.

Keywords Spectrofluorimetry · Chemometrics · Naproxen · Salicylic acid · Salicyluric acid · β-Cyclodextrin

Abbreviations *EEFMs* excitation-emission fluorescence matrices · *NX* naproxen · *SA* salicylic acid · *SU* salicyluric acid · *GRAM* generalized rank annihilation method · *IRGRAM* iteratively reweighted generalized rank annihilation method · *PARAFAC* parallel factor analysis · *SWATLD* self-weighted alternating trilinear decomposition

J.A. Arancibia · A.C. Olivieri · G.M. Escandar (✉)
Departamento de Química Analítica,
Facultad de Ciencias Bioquímicas y Farmacéuticas,
Universidad Nacional de Rosario,
Suipacha 531 (S2002LRK) Rosario, Argentina
e-mail: gescanda@fbioyf.unr.edu.ar

Introduction

In recent years, multivariate calibration methods have been gaining importance in the resolution of complex multicomponent mixtures [1]. First-order methods such as partial least-squares (PLS) regression employ vectorial data (i.e., spectra collected at a number of wavelengths). They require the construction of large training sets for a successful calibration before prediction is made, and are sensitive to the presence of unmodelled components, as is frequently observed in the analysis of biological samples. On the other hand, matrix data such as excitation-emission fluorescence matrices (EEFMs) can be adequately analyzed by second-order methods. They allow spectra and relative concentrations of individual mixture components to be extracted directly, in the presence of any number of uncalibrated constituents. This can be achieved using only a very small sample set, as will be illustrated in the present work.

Experimental applications of EEFMs that have been reported so far are rather scarce, and have been generally restricted to mixtures of known components such as polycyclic aromatic hydrocarbons [2, 3]. To the best of our knowledge, this type of analysis has not been previously performed on biological samples.

In this report we selected, as convenient examples, the simultaneous determination of the anti-inflammatories naproxen [NX: (+)-2-(6-methoxy-2-naphtyl)propionic acid] and salicylic acid (SA: 2-hydroxybenzoic acid) in human serum and of NX, SA, and the main metabolite of the latter, salicyluric acid (SU: 2-hydroxybenzoylaminoacetate) [4] in human urine. The fluorescent properties of these three compounds are well known, and some previous knowledge on the application of first-order calibration (at the PLS level) to serum samples is already available from our laboratory [5]. Previous non-chromatographic analyses of the presently studied analytes have been reported in the literature [6, 7, 8, 9, 10, 11]. It should be noticed, however, that the simultaneous spectrofluorimetric resolution of NX, SA, and SU in urine has not been previously reported.

The objective of this work was to compare the relative advantages and disadvantages of a first-order (PLS) and

three different second-order fluorescence methods [12, 13, 14, 15] for the analysis of biological samples, both from the experimental point of view and from the quality of the analytical results. The performance of the methods is shown to improve through the formation of inclusion complexes between β-cyclodextrin (β-CD) and the investigated compounds. These complexes may show enhanced fluorescent properties with useful analytical implications [16, 17, 18, 19]. Whereas the NX-β-CD and SA-β-CD complex formations have been previously reported [20, 21], the corresponding analysis of the SU-β-CD host-guest complex system is assessed in the present work.

Experimental

Reagents

Analytical-reagent grade chemicals were employed in all experiments. Stock solutions (all 500 μg mL^{-1}) of sodium naproxenate (Sigma, St. Louis, MO, USA), sodium salicylate (Merck, Darmstadt, Germany) and salicyluric acid (Aldrich, Milwaukee, WI, USA) were prepared by dissolving each compound in doubly distilled water. From these concentrated solutions, either 1 or 5 μg mL^{-1} working solutions were prepared. A buffer solution of borax of pH 9.0 was prepared in a 100 ml volumetric flask by adding 4.6 ml of 0.10 M HCl to 50 ml of 0.025 M sodium tetraborate decahydrate (Analar, London, UK) and diluting to the mark with water. β-CD (Aldrich, Milwaukee, WI, USA) was recrystallized twice from water, and 0.01 M solutions were prepared either in water or in borax buffer. Sodium dodecyl sulfate (SDS) was obtained from Merck (Darmstadt, Germany) and α-CD from Cyclolab (Budapest, Hungary). Both were used as received. Aqueous 0.01 M solutions of the latter two compounds were prepared.

Apparatus

Synchronous fluorescence spectra were measured on a Shimadzu RF 5300 PC spectrofluorimeter, equipped with a 150 W xenon lamp, and using 1.00 cm quartz cells. They were collected each 1 nm in the range 325–500 nm (using $\Delta\lambda$=105 nm), converted to, and saved in, ASCII format, and transferred to a Pentium III microcomputer. Excitation-emission fluorescence matrices were measured on an Aminco Bowman (Urbana, IL, USA) Series 2 spectrofluorimeter fitted with a 150 W continuous xenon lamp and interfaced to a 486 personal computer. Data acquisition and data analysis were performed with AB2 software, running under OS/2 version 4.0. Fluorescence measurements both for serum and urine were carried out using 1.00 cm quartz cells, slit widths, 4 nm, wavelength excitation range, 200–350 nm (each 5 nm) and wavelength emission range, 330–460 nm (each 1 nm).

Influence of β-cyclodextrin concentration

In the inclusion constant determination for SU-β-CD complex, the concentration of SU was held constant at 6.4×10^{-6} M while the β-CD was varied from 0 to 0.01 M. The experiment was performed in NaOH solution (pH 9) and in duplicate. The fluorescence emission was measured at λ_{exc}=290 nm and λ_{em}=410 nm.

Serum samples

The construction of a calibration set for applying PLS to serum samples has already been described in the literature [5]. No such training set is needed in the present case, where standard addition is applied to second-order fluorescence data. Accordingly, only a validation set of nine samples was prepared by spiking different human sera with NX and SA in order to obtain the five concentrations levels provided by a central composite design, and their corresponding partners with addition of standard. A volume of 100 μL of each serum sample was diluted to 1 mL with water. 10 μL of this solution were placed in a 2 mL flask and the corresponding volumes of both NX and SA stock solutions were added. Completion to the mark was achieved with 0.01 M β-cyclodextrin aqueous solution. The highest concentrations of both NX and SA in the flasks were 125 and 250 ng mL^{-1}, respectively. These concentrations correspond to original values in serum of 250 and 500 mg L^{-1} for NX and SA, respectively, and are the maximum concentrations found under normal therapeutic doses. Since SA emits fluorescence at pH>4 (when present as salicylate) and the fluorescence emission of NX remains constant at a wide pH range (2–12) [20], the working pH was that obtained by the dissolution of the compounds (almost neutral). Each serum sample was prepared in duplicate and no pre-treatment was required. The EEFMs were then read for these samples, and for additional ones prepared in the same manner as described above, but containing final concentrations which were, for both analytes, 50, 100, and 150 ng mL^{-1} higher than those given by the central composite design. The obtained data were subjected to second-order data analysis, as described below.

Urine samples

For the first-order PLS model building, a calibration set of 30 different urine samples was prepared with the five concentration levels of NX, SA and SU provided by two identical 15-sample central composite designs. The levels correspond to values in the range 0–200 ng mL^{-1} for NX and 0–300 ng mL^{-1} for both SA and SU. The experiments were performed in both the absence and the presence of β-CD at a constant concentration of 0.01 M, and at pH 9 (borax buffer). At this pH value, the three compounds emit fluorescence [8, 20] while the deprotonation of the β-CD, which would occur at high pH, is avoided. The validation experiment was carried out using 12 urine samples different to those employed for calibration, with random concentrations of NX, SA and SU (all lying in the corresponding calibration range). Each sample was prepared in triplicate. The synchronous fluorescence spectra were then read (in random order with respect to analytes concentrations) and subjected to PLS analysis.

For the second-order data analysis, a set of 12 validation samples was prepared in the same form and with the same concentrations of NX, SA and SU as those employed for validating the PLS model. The EEFMs were read for each of these samples, and for additional ones prepared in a similar way but having final concentrations of NX, SA and SU 200 ng mL^{-1} higher.

PLS analysis

The theory of PLS is well known [22]. Briefly, the method involves a calibration step in which the relation between spectra and analyte concentrations is estimated from a set of reference samples, and a prediction step in which the results of the calibration are used to estimate the component concentrations in unknown samples. PLS was applied with the program MULTIVAR, written in Visual Basic 5.0, and freely available on the internet at ftp://www.fbioyf.unr.edu.ar/cientifico/multivar.exe. The program is based on a previously described algorithm [23, 24].

Theory of second-order methods

Second-order fluorescence data can be analyzed by two types of methods: 1) those based on solving an eigenvalue problem, such as the generalized rank annihilation method (GRAM) [12] and its iteratively reweighted variant (IRGRAM) [25] and 2) those employing a least-squares minimization, as implemented in both parallel factor analysis (PARAFAC) [14] and self-weighted alternating trilinear decomposition (SWATLD) [15]. EEFMs for a group of I samples can be represented as:

$$F_{ijl} = \sum_{k=1}^{K} c_{ik}\, X_{kl}\, Y_{kj} \qquad (1)$$

where F_{ijl} is an element of the cube F (dimensions=$I \times J \times L$, I=number of samples, J=number of excitation wavelengths, L=number of emission wavelengths), i.e., the fluorescence intensity for sample i at the emission wavelength j and excitation wavelength l, c_{ik} is the concentration of k in sample i, X_{kl} is the excitation profile of analyte k at excitation wavelength l, Y_{kj} is the emission profile of analyte k at wavelength j and K is the total number of analytes (here and elsewhere, error terms are dropped from the equations in order to simplify the notation). Equation 1 shows that EEFMs follow a trilinear model to which programs such as GRAM, PARAFAC, and SWATLD can be successfully applied. For specific details see the Appendix.

Although ambiguities remain in the absolute component concentrations and chemical identities when these methods are applied, they can be solved by: a) using samples of known concentration and b) comparing the individual excitation or emission profiles with those of standards. In the present report, IRGRAM was implemented with an in-house MATLAB 5.3 routine [26] based in ref. [25], and PARAFAC with the MATLAB routine developed by Bro [27] and available from the internet at http://www.models.kvl.dk/. The routine for applying SWATLD was generously supplied by NM Faber.

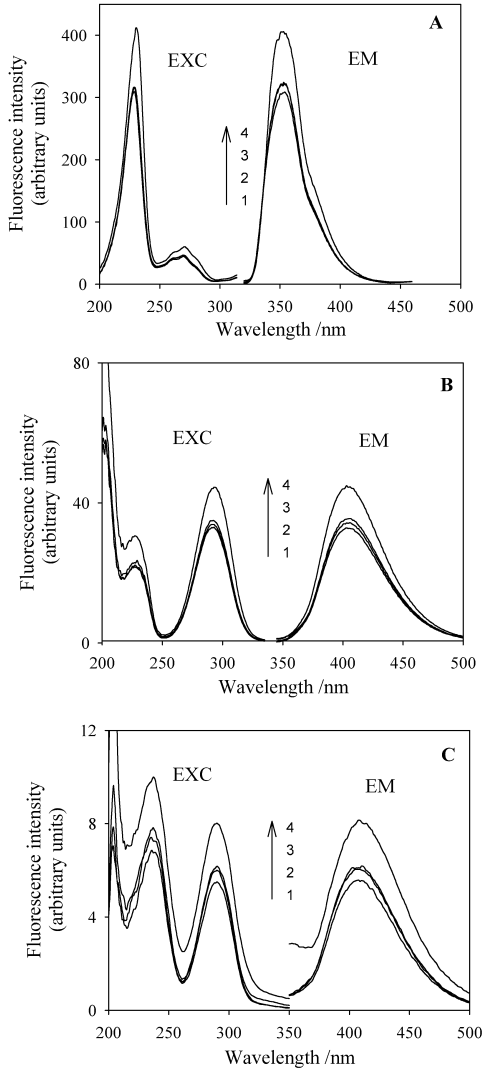

Fig. 1A–C Excitation and emission spectra of **A** naproxen (250 ng mL^{-1}), **B** salicylic acid (250 ng mL^{-1}) and **C** salicyluric acid (250 ng mL^{-1}). In **A** and **B**, (*1*) aqueous solution, (*2*) 0.01 M SDS micellar solution (*3*) 0.01 M α-cyclodextrin, and (*4*) 0.01 M β-cyclodextrin, whereas in **C**, (*1*) 0.01 M SDS micellar solution, (*2*) aqueous solution, (*3*) 0.01 M α-cyclodextrin and (*4*) 0.01 M β-cyclodextrin

Results and discussion

With the purpose of increasing the native fluorescence intensities of NX, SA, and SU, and thus improving the performance of the analytical methods, different organized media such as micelles and cyclodextrins (namely, SDS, α-CD, and β-CD) were checked in the corresponding optimum working conditions. Among the investigated systems, β-CD produced a significant enhancement of the fluorescence spectra of the analytes (Fig. 1), while not affecting those of serum and urine, and was therefore selected for subsequent experiments. The relative interplay of spectra for analytes and background components is such that the presence of β-CD is critical for the success of the determinations in urine, whereas in serum the performances in the absence and in the presence of β-CD are similar.

In previous reports we have demonstrated that both NX and SA form 1:1 inclusion complexes with β-CD [20, 21], with equilibrium constants which are 2100 and 516 M^{-1} for the naproxen complexes at pH 2.5 and 10, respectively, and 100 M^{-1} for the salicylate complex at neutral pH. In the present paper, the inclusion constant value for the SU-β-CD complex was fluorimetrically calculated. The obtained value (210±12 at pH=9.0) suggests only a moderate host-guest interaction.

Serum analysis

Figure 2 shows the excitation-emission fluorescence contour plots corresponding to a typical serum and to aqueous solutions containing NX and SA at neutral pH. As can be

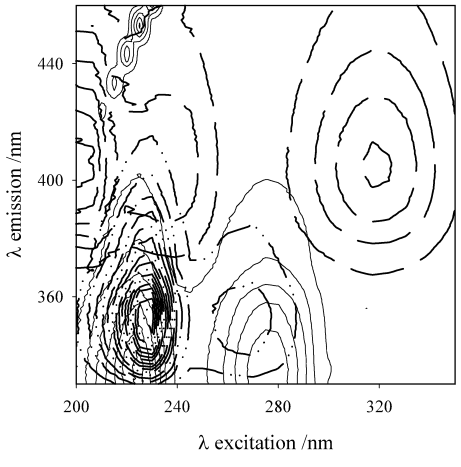

Fig. 2 Contour plots corresponding to the fluorescence emission-excitation matrices for: (– – –) 250 ng mL^{-1} aqueous salicylate, (– ·· –) 100 ng mL^{-1} aqueous naproxen and (—) a typical human serum

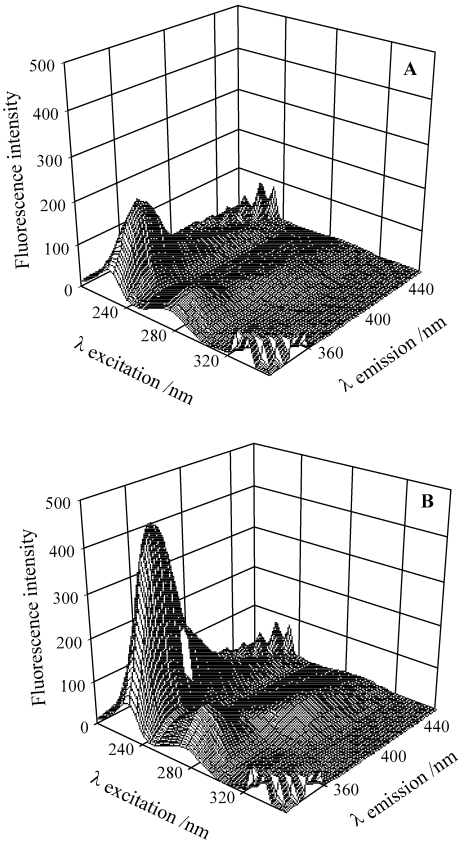

Fig. 3 **A** Three-dimensional plot of the EEFM for a typical validation serum sample containing NX (125 ng mL^{-1}) and SA (130 ng mL^{-1}). **B** Same as **A** after the standard addition of 100 ng mL^{-1} of NX and SA

observed, the overlapping of the spectra for serum and the investigated compounds is significant and hinders their direct fluorimetric determination. In principle, this problem can be overcome by using first-order data (fluorescence emission collected in the range 340–500 nm at an excitation wavelength of 315 nm) in combination with PLS [5]. This requires the construction of a suitable calibration set (formed by 25 samples according to ref. [5]), containing both analytes and sera from different individuals. However, the use of EEFMs seems to be a simpler experimental approach, as described below. In what follows, serum results in the presence of β-CD will be discussed, in order to allow for a final comparison with those obtained in urine under the best experimental conditions. Nevertheless, similar figures of merit are obtained in serum in the absence of β-CD.

Examples of EEFMs for a typical serum, before and after standard addition, are provided in Fig. 3. The latter procedure was adopted because it takes into account possible spectral variations of the analytes due to chemical interactions with a matrix of the complexity of human serum. Before applying the chemometric models for the EEFMs, a selection of wavelength ranges appropriate for each compound was made. These ranges were restricted upon suitable consideration of the contour plots shown in Fig. 2: 200–300 nm (excitation) and 330–370 nm (emission) for NX, and 285–340 (excitation) and 380–440 (emission) for SA.

Table 1 shows the nominal and predicted analyte concentrations and the corresponding recoveries using IRGRAM, PARAFAC, and SWATLD, applying both single and multiple standard addition, and using two or three

Table 1 Serum recovery results using second-order data

	Actual values (ng mL^{-1})	Single calibration sample						Multiple (3) calibration samples					
		IRGRAM		PARAFAC		SWATLD		IRGRAM		PARAFAC		SWATLD	
		Found[a] (ng mL^{-1})	Recovery (%)	Found[a] (ng mL^{-1})	Recovery (%)	Found[a] (ng mL^{-1})	Recovery (%)	Found[a] (ng mL^{-1})	Recovery (%)	Found[a] (ng mL^{-1})	Recovery (%)	Found[a] (ng mL^{-1})	Recovery (%)
NX	10	7 (1)	70	6 (1)	60	6 (1)	60	7 (1)	70	8 (1)	80	9 (1)	90
SA	130	167 (1)	128	158 (1)	122	147 (1)	113	164 (1)	126	140 (1)	108	137 (1)	105
NX	125	108 (5)	86	108 (1)	86	107 (1)	86	112 (5)	90	123 (1)	98	123 (1)	98
SA	130	126 (1)	97	129 (1)	103	132 (1)	101	129 (1)	99	135 (1)	104	132 (1)	101
NX	65	65 (3)	100	65 (1)	100	65 (1)	99	66 (2)	102	66 (1)	102	66 (1)	101
SA	20	16 (1)	80	15 (1)	77	15 (1)	74	16 (1)	80	20 (1)	101	20 (1)	102
NX	65	67 (5)	103	66 (1)	102	66 (1)	101	64 (5)	98	62 (1)	95	62 (1)	95
SA	250	296 (5)	118	269 (3)	108	221 (1)	88	304 (5)	121	282 (3)	113	236 (1)	94
NX	27	24 (3)	89	23 (1)	84	23 (1)	83	23 (3)	85	26 (1)	95	26 (1)	95
SA	54	56 (1)	103	56 (1)	104	56 (1)	104	59 (1)	110	51 (1)	94	51 (1)	95
NX	108	108 (4)	100	108 (1)	100	107 (1)	99	109 (4)	101	106 (1)	98	105 (1)	98
SA	54	57 (1)	105	57 (1)	105	56 (1)	104	58 (1)	107	59 (1)	109	58 (1)	108
NX	27	26 (4)	96	25 (1)	94	25 (1)	94	25 (4)	93	25 (1)	91	25 (1)	91
SA	217	253 (4)	117	219 (1)	101	198 (1)	91	257 (4)	118	251 (2)	115	215 (1)	99
NX	108	103 (5)	95	104 (1)	96	103 (1)	96	105 (5)	97	102 (1)	95	103 (1)	95
SA	217	236 (2)	109	234 (2)	108	211 (1)	97	234 (2)	108	238 (1)	110	217 (1)	100
NX	65	63 (3)	97	62 (1)	95	62 (1)	95	63 (4)	97	69 (1)	106	69 (1)	105
SA	130	148 (1)	114	148 (1)	114	148 (1)	114	140 (1)	108	141 (1)	108	134 (1)	103

[a] Standard deviation (in parenthesis) refer to random errors, and were calculated according to ref. [31]

Table 2 Validation statistical results for the simultaneous determination of NX and SA in serum by second-order data

	IRGRAM		PARAFAC		SWATLD	
Single standard addition						
	RMSEP (ng mL^{-1})	R^2	RMSEP (ng mL^{-1})	R^2	RMSEP (ng mL^{-1})	R^2
NX	7	0.974	7	0.973	7	0.970
SA	26	0.896	15	0.967	15	0.964
Multiple (3) standard additions						
NX	5	0.984	3	0.994	3	0.995
SA	27	0.885	19	0.945	6	0.995

[a]RMSEP (root mean square errors of prediction)= $\left[\frac{1}{I-1}\sum_{1}^{I}\left(c_{act}-c_{pred}\right)^2\right]^{1/2}$ where I is the number of prediction samples, c_{act} and c_{pred} are the actual and predicted concentrations respectively. R^2 (square correlation coefficient)= $-\frac{\sum_{1}^{I}(c_{act}-c_{pred})^2}{\sum_{1}^{I}(c_{act}-c)^2}$

factors depending on the specific sample (they were selected by trial and error, see Appendix). An inspection of the values quoted in Table 1 for NX using either standard addition mode seems to indicate reasonably good figures with all three computational methods, although multiple addition yields better results. In all cases the standard deviations in the predicted concentrations are in the order of 1–5 ng mL^{-1}, consistent with error propagation during sample preparation. Concerning SA, an interesting trend is detected at high concentrations: only the values predicted for this analyte by SWATLD are in agreement with the nominal ones. A likely explanation is that this approach is less susceptible to slight deviations from linearity. A summary of validation statistics for both analytes is shown in Table 2, which further demonstrates that, in general, the procedure where several standard additions are used shows a better analytical performance.

In order to gain further insight into the accuracy of the methods, linear regression analysis of nominal versus found concentration values was applied. The estimated intercept and slope ($\hat{a}$ and $\hat{b}$ respectively) were compared with their ideal values of 0 and 1 using the elliptical joint confidence region (EJCR) test. The boundary of the ellipse is determined by the magnitude of experimental errors and by the degrees of confidence chosen, and is described by the following equation [28]:

$$N_{dat}(a-\hat{a})^2 + 2(\sum c_{act})(a-\hat{a})(b-\hat{b}) + (\sum c_{act})^2(b-\hat{b})^2 = 2s^2F_{\alpha,2,Ndat-2} \quad (2)$$

where N_{dat} is the number of data points, c_{act} are the actual concentrations, s^2 is the regression variance and $F_{\alpha,2,Ndat-2}$ is the critical F value with 2 and (N_{dat}–2) degrees of freedom at a given 100×(1–α) confidence level, usually 95%. If the point (a=0, b=1) is inside the EJCR, it can be concluded that constant and proportional bias are absent. For multicomponent mixtures, results should be studied for all analytes simultaneously in Eq. 2, rather than performing indi-

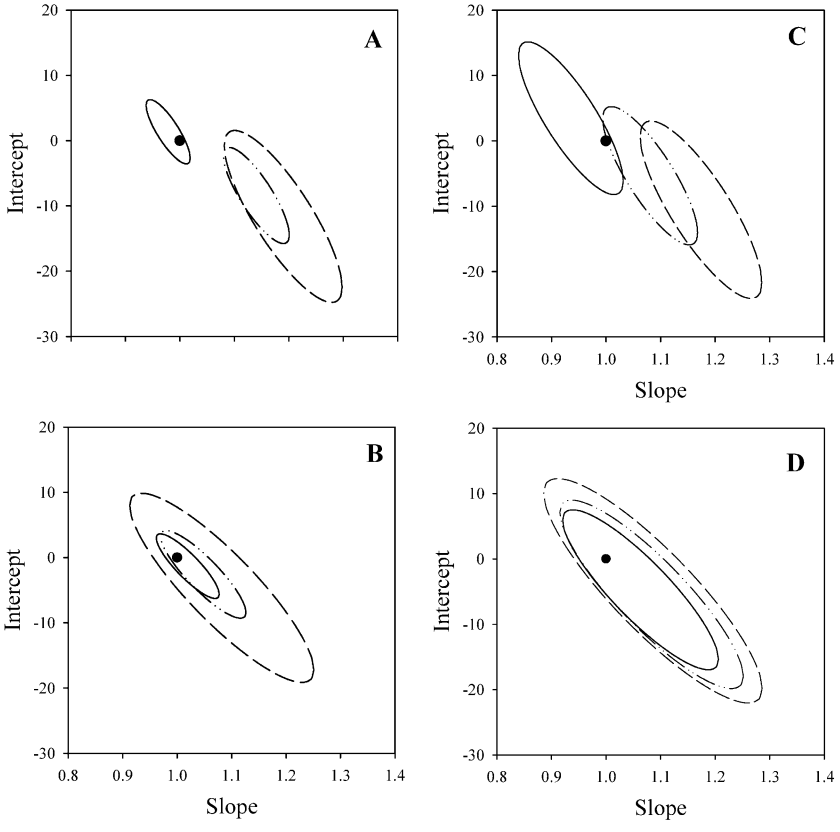

Fig. 4A–D Elliptical joint confidence regions for the slope and intercept corresponding to regressions of predicted vs. actual concentrations of salicylate and naproxen applying the methods (– – –) IRGRAM, (– ·· –) PARAFAC and (—) SWATLD. **A** Multiple calibration samples using all data. **B** Multiple calibration samples excluding the highest concentration values of salicylate (see text). **C** Two calibration samples using all data. **D** Two calibration samples excluding the highest concentration values of salicylate. The *black circles* mark the theoretical (a=0, b=1) points

vidual tests for each component [28]. Figure 4A shows the EJCR plots for the three employed chemometric methods when multiple additions are performed, using the 18 concentration values corresponding to the two analytes in the nine sample set. As can be seen, only the EJCR for SWATLD results contains the theoretical (a=0, b=1) point. If the test is carried out without considering the higher values of SA (specifically, 217 and 250 ng mL^{-1}) the three ellipses contain the ideal points (Fig. 4B), but that corresponding to IRGRAM shows a rather large (and undesirable) size.

The EJCR tests corresponding to the experiments conducted under single standard addition are displayed in Figs. 4C and 4D. The observed behavior is similar to that found for multiple additions, but all ellipse sizes are significantly larger. The conclusion is that this procedure, although more rapid and simple, is, however, less precise.

Urine analysis

First-order data

Synchronous fluorescence spectra were used for analyzing urine containing NX, SA, and SU. This instrumental technique usually provides better resolution than either emission or excitation spectra. The first step for its successful application is the estimation of the optimum value of Δλ [29], which can be done by resorting to excitation-emission contour plots for overlapping signals. Fig. 5A shows the plot corresponding to a typical urine and Fig. 5B to aqueous solutions containing NX, SA, or SU at pH= 9.0. As can be appreciated in these figures, human urine shows a broad fluorescence spectrum, which partially overlaps with those of the investigated analytes. It was found that Δλ=105 nm provides the best differentiation (the corresponding straight line in Fig. 5B shows the path for the selected Δλ). Other choices of Δλ led to poorer results. Figure 6 shows the resulting synchronous fluorescence spectra for urine and the three studied analytes in the absence and in the presence of β-CD. While the urine signal remains almost constant, the analytes fluorescence signals are intensified by β-CD, as expected from Fig. 1.

Before calibration with the PLS method, the most appropriate wavelength ranges for quantitating each analyte were selected. This was done by calculating the optimum cross-validation variance for different wavelength windows [30] and selecting those leading to a minimum variance for each analyte. According to Table 3, the optimum ranges include the major synchronous fluorescence emission peaks observed in Fig. 6. Table 3 also shows the number of factors (estimated from cross-validation) and the calibration and validation statistics, while Table 4 shows the recoveries obtained for the studied mixtures. It can be concluded that the technique yields significantly better results and statistical indicators when the experiment is run in the presence of β-CD. The EJCR test was also applied using the 36 concentration values corresponding to the three analytes in the 12-sample validation set. In agreement

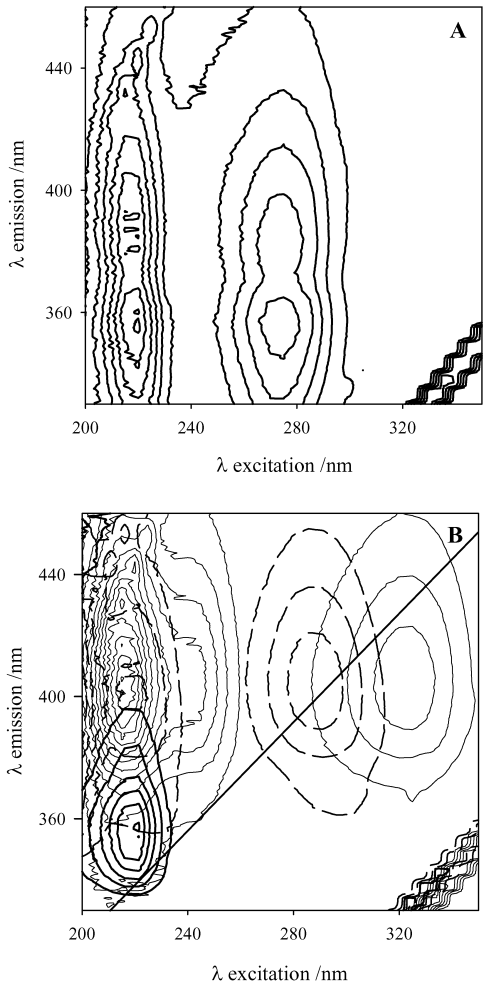

Fig. 5A,B Contour plots of the total fluorescence spectra of: **A** a typical human urine, **B** (—), salicylic acid (—) and salicyluric acid (—). The *straight solid line* represents the synchronous path using Δλ=105 nm. C_{NX}=55 ng mL^{-1}, C_{SA}=206 ng mL^{-1}, C_{SU}=194 ng mL^{-1}, pH=9.0

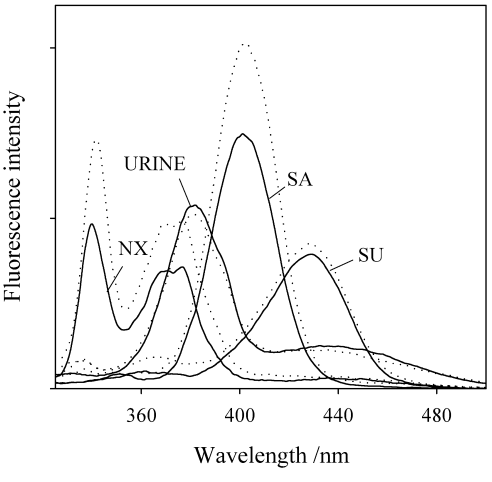

Fig. 6 Synchronous spectra of naproxenate (*NX*), salicylic acid (*SA*), salicyluric acid (*SU*) and a typical human urine in the absence (—) and in the presence (· · · ·) of β-cyclodextrin. C_{NX}= 110 ng mL^{-1}, C_{SA}=206 ng mL^{-1}, C_{SU}=194 ng mL^{-1}, $C_{\beta\text{-CD}}$=9.9× 10^{-3} M, Δλ=105 nm, pH=9.0

Table 3 Calibration and validation statistical results for the simultaneous determination of NX, SA and SU in urine by synchronous fluorescence combined with PLS[a]

Calibration statistical parameters												
	Without β-CD						With β-CD					
	Wavelength range (nm)	*A*	RMSEC (ng mL^{-1})	R^2	SEN[b]	SEL[b]	Wavelength range (nm)	*A*	RMSEC (ng mL^{-1})	R^2	SEN	SEL
NX	375–465	4	15	0.934	0.16	0.24	325–415	6	9	0.975	0.22	0.25
SA	385–425	5	7	0.993	0.19	0.26	355–455	6	7	0.994	0.26	0.29
SU	395–465	5	16	0.965	0.11	0.18	415–455	6	12	0.980	0.13	0.19
Validation statistical parameters												
	Without β-CD						With β-CD					
			RMSEP	R^2					RMSEP	R^2		
NX			21	0.882					10	0.976		
SA			18	0.947					11	0.981		
SU			20	0.950					15	0.972		

[a]*A*=Number of factors. RMSEC (root mean square errors of calibration)= $\left[\frac{1}{I-1}\sum_{1}^{I}(c_{act}-c_{pred})^2\right]^{1/2}$ where *I* is the number of calibration samples, RMSEP and R^2 as in Table 2. SEN and SEL are the sensitivity and selectivity respectively.
[b]These parameters were calculated in the full spectral range 325–500 nm

with Table 4, the EJCR for the experiments in the presence of β-CD contains the theoretical (a=0, b=1) point, in contrast with the case where the β-CD is absent.

PLS sensitivity and selectivity were also calculated. For a proper comparison of these parameters when β-CD is added, the same spectral range should be used for all analytes. Table 3 shows the corresponding results, which reveal some interesting details. The sensitivity increases with the addition of β-CD, with the relative increase being larger for NX and SA, as expected from the inspection of Fig. 6, which qualitatively shows that the effect of β-CD on the fluorescence signals of NX and SA are larger than that on SU. The selectivity, on the other hand, is not seriously affected by the presence of β-CD (Table 3), which stems from the fact that the shapes of the spectra do not significantly change upon inclusion of the cyclodextrin. The conclusion from Table 3 and the EJCR analysis is that the addition of cyclodextrin not only increases the sensitivity of the determination, but also produces more accurate results.

Second-order data

Second-order EEFMs were also produced for the validation set of samples studied by the PLS method. A selection of the excitation and emission wavelength ranges for each analyte was made upon suitable consideration of the contour plots shown in Fig. 5: 200–300 nm (excitation) and 330–410 nm (emission) for NX, 275–325 nm (excitation) and 360–430 nm (emission) for SA, and 300–350 nm (excitation) and 370–460 nm (emission) for SU. Within each of these selected regions, the number of factors to be used by second-order methods was estimated as three, by trial and error.

Single standard addition was first applied, employing βCD to improve the sensitivity, with the result that IRGRAM, PARAFAC, and SWATLD models gave comparably acceptable results for NX, i.e., RMSEP=14 ng mL^{-1}, REP=14%, and R^2=0.950. The corresponding analyses of SA and SU rendered, in contrast to NX, discouragingly poor recoveries. We believe this may be ascribed to a lack of selectivity, i.e., to the fact that the fluorescence emission profiles for SA and SU are almost identical, precluding the successful decomposition of the second-order data for these particular analytes. This can be qualitatively appreciated in Fig. 5B, where the projections of the contour plots of SA and SU onto the emission axis are seen to be very similar. From a mathematical point of view, if two analytes are present with a common emission profile Y_j, the fluorescence cube elements in *I* samples are given by (see Eq. 1):

$$F_{ijl} = Y_j \sum_{k=1}^{2} c_{ik}\, X_{kl} \qquad (3)$$

where Y_j is now an analyte-independent proportional constant. Eq. 3 indicates that the data lose the trilinear property, and the cube decomposition is no longer unique. In view of the results commented above, no efforts were made to apply multiple standard addition.

Comparison of methods

In comparing the performances of the investigated multivariate techniques, one may note that the best second-order results for simultaneously determining NX and SA in serum (RMSEP=3 ng mL^{-1} for NX and 6 ng mL^{-1} for SA, obtained using SWATLD for data processing, see Table 2)

Table 4 Urine recovery results using first-order data (PLS)

	Actual values (ng mL^{-1})	Without β-CD		With β-CD	
		Found values[a] (ng mL^{-1})	Recovery (%)	Found values[a] (ng mL^{-1})	Recovery (%)
NX	152	153 (6)	101	167 (6)	110
SA	61	54 (4)	89	63 (4)	104
SU	265	268 (7)	101	239 (9)	90
NX	123	118 (7)	96	128 (7)	104
SA	72	66 (4)	92	78.5 (4)	109
SU	239	249 (7)	105	227 (10)	95
NX	59	47 (7)	79	60.2 (7)	101
SA	174	154 (4)	89	161 (4)	93
SU	274	275 (8)	100	252 (10)	92
NX	0	–3	-	–8	-
SA	261	255 (7)	98	268 (4)	103
SU	151	142 (8)	94	135 (10)	89
NX	192	140 (10)	73	172 (9)	90
SA	291	282 (8)	97	293 (7)	101
SU	184	176 (9)	96	173 (9)	94
NX	29	10 (8)	36	33 (6)	113
SA	157	146 (5)	93	171 (5)	109
SU	276	270 (7)	98	272 (10)	99
NX	183	160 (20)	87	177 (10)	97
SA	145	179 (10)	123	146 (10)	101
SU	117	73 (10)	62	141(10)	120
NX	146	200 (10)	137	159 (10)	109
SA	272	268 (8)	99	266 (5)	98
SU	149	127 (8)	85	151 (10)	101
NX	55	60 (10)	120	59 (10)	108
SA	206	187 (5)	91	218 (6)	106
SU	194	161 (6)	83	183 (9)	94
NX	100	130 (40)	130	105 (9)	105
SA	120	141 (8)	118	136 (5)	113
SU	45	34 (7)	75	41 (8)	92
NX	42	70 (10)	166	43 (7)	103
SA	96	93 (4)	97	103 (5)	107
SU	29	33 (7)	116	32 (7)	112
NX	80	89 (40)	111	82 (7)	103
SA	190	218 (8)	115	207 (4)	109
SU	60	42 (8)	70	55 (7)	91

[a]Standard deviations (in parenthesis) refer to random errors, and were calculated according to ref. [32]

are as satisfactory as those furnished by PLS analysis (RMSEP=3 ng mL^{-1} for both NX and SA [5]). As to the determination of NX in urine, the corresponding second-order result (RMSEP=14 ng mL^{-1}) is slightly worse than that obtained from PLS (RMSEP=10 ng mL^{-1}), but may be considered acceptable in view of the complexity of the samples at hand. A cross-comparison of the results in both biological fluids indicates that the analytical performance is poorer in urine as compared to serum, due to the strong spectral overlapping and increased number of interfering components present in the former.

From the experimental point of view, the use of second-order data is considerably simpler to implement than the PLS calibration. There is no need for measuring spectra of a large number of calibration samples, only for obtaining the EEFMs for the unknown sample, with and without the addition of standards of known concentration. Specific software needing a minimum of MATLAB programming skill is freely available on the internet. The examples analyzed in the present paper illustrate a possible limitation in the applicability of the technique, namely the occurrence of an adequate selectivity towards specific analytes.

The final conclusion is that (excluding the cases of lack of trilinearity) the methods that use second-order data are an excellent tool for the determination of compounds in complex matrices, due to their simplicity and analytical performance.

Acknowledgements This research was supported by the Universidad Nacional de Rosario, the Agencia Nacional de Promoción Científica y Tecnológica (Project PICT99 No. 06–06078), and the Consejo Nacional de Investigaciones Científicas y Técnicas (Project PIP No. 431). A.C.O. is a fellow of the John Simon Guggenheim Memorial Foundation (2001–2002).

Appendix

The GRAM model, as implemented in the standard addition version [12], employs two matrices: M (the $J\times L$ EEFM of a given unknown sample) and W (the $J\times L$ EEFM of the unknown sample to which a known concentration of standard for component k has been added). In principle, W can be written as:

$$\mathbf{W}=\mathbf{M}+c_{k,\text{added}}\mathbf{Q}_k \tag{A1}$$

where $\mathbf{Q}_k$ is the EEFM for analyte k at unit concentration and $c_{k,\text{added}}$ is the added concentration. The first step in GRAM is the singular value decomposition of W:

$$\mathbf{W}=\mathbf{U}\,\mathbf{S}\,\mathbf{V}^{\mathrm{T}} \tag{A2}$$

In Eq. A2, only a number of significant components is used, corresponding, in the ideal case, to the number of fluorescent sample constituents. In complex biological samples, complete knowledge of the chemical components is lacking, and the number of factors is usually established by trial and error or by principal component analysis of M.

The next step is the projection of M onto the principal components of W, yielding the following eigenvalue-eigenvector problem [12]:

$$\mathbf{U}^{\mathrm{T}}\,\mathbf{M}\mathbf{V}\,\mathbf{S}^{-1}\,\mathbf{Z}=\lambda\mathbf{Z} \tag{A3}$$

where:

$$\mathbf{Z}=\mathbf{S}\,\mathbf{V}^{\mathrm{T}}\,\mathbf{Y}^{\mathrm{T}+} \tag{A4}$$

In Eq. A4, Y contains the fluorescence emission profiles, T and T+ indicate transposition and pseudoinverse respectively, and λ is a $k\times k$ matrix whose diagonal elements are the individual eigenvalues λ_k. It can be shown that each λ_k corresponds to the following ratio of concentrations:

$$\lambda_k = c_{k,\mathrm{M}}/c_{k,\mathrm{W}} = c_{k,\mathrm{M}}/(c_{k,\mathrm{M}} + c_{k,\text{added}}) \tag{A5}$$

where $c_{k,\mathrm{M}}$ and $c_{k,\mathrm{W}}$ are the concentrations of component k in the samples whose EEFMs are M and W, respectively. Therefore, prediction proceeds as:

$$c_{k,\mathrm{M}} = \lambda_k c_{k,\mathrm{added}}/(1-\lambda_k) \tag{A6}$$

whereas for all other components present in the unknown sample, $c_{k,\mathrm{added}}$=0 and λ_k=1. Originally, GRAM was formulated in terms of a single calibration sample, but can also be adapted to accommodate multiple calibration samples [31], using the following equation:

$$\mathbf{W}^{(n)} = \mathbf{M} + \mathbf{Q}_k^{(n)} = \mathbf{M} + c_{k,added}^{(n)}\mathbf{Q}_k \tag{A7}$$

where $\mathrm{W}^{(n)}$ represents each of the matrices corresponding to the *n*th. addition of a standard of component *k*, and $\mathbf{Q}_k^{(n)}$ is the EEFM for the analyte at a concentration corresponding to the *n*th. addition ($c_{k,added}^{(n)}$). Averaging over the *N* standard additions leads to:

$$\mathbf{W} = \frac{\sum_{n=1}^{N}\mathbf{W}^{(n)}}{N} = \mathbf{M} + \frac{\sum_{n=1}^{N} c_{k,added}^{(n)}}{N}\mathbf{Q}_k = \mathbf{M} + \bar{c}_{k,added}\mathbf{Q}_k \tag{A8}$$

where $\bar{c}_{k,added}$ is the mean added concentration. Equation A8 is analogous to Eq. A1, and thus prediction proceeds in this case as:

$$c_{k,\mathrm{M}} = \lambda_k\, \bar{c}_{k,added}/(1-\lambda_k) \tag{A9}$$

The final step is to obtain the individual excitation and emission profiles, which assist in matching the eigenvalues with the correct sample components. They are available from:

$$\mathbf{X} = (c_{k,\mathrm{M}})^{-1}\mathbf{U}\,\mathbf{Z} \tag{A10}$$

$$\mathbf{Y}^{\mathrm{T}} = (\mathbf{V}\,\mathbf{S}^{-1}\,\mathbf{Z})^{+} \tag{A11}$$

In a variant of GRAM, recently developed by Faber et. al. [13], the matrices M and W are first linearly combined to give an "effective" $\mathrm{W}_{\mathrm{eff}}$ matrix:

$$\mathbf{W}_{\mathrm{eff}} = \alpha\,\mathbf{W} + (1-\alpha)\mathbf{M} \tag{A12}$$

where α is a coefficient which can be found iteratively. After calculating the matrix $\mathrm{W}_{\mathrm{eff}}$ starting with α=1, Eqs. A2 and A3 are applied, with W being replaced by $\mathrm{W}_{\mathrm{eff}}$. It can be easily shown that the predicted concentration of analyte *k* is given in this case as:

$$c_{k,\mathrm{M}} = \alpha\,\lambda_k\,\bar{c}_{k,added}/(1-\lambda_k) \tag{A13}$$

Once this step is completed, a new value of α is computed as the ratio $\bar{c}_{k,added}/\,c_{k,\mathrm{M}}$, and Eqs. A12 and A13 are successively applied until α converges. This procedure has been shown to improve the intrinsic prediction bias of the original GRAM [13].

In the least-squares method known as PARAFAC, the cube of EEFMs is fitted by minimizing the following objective function:

$$S(\mathbf{X},\mathbf{Y},\mathbf{c}_i) = \sum_{i=1}^{I}\left\|\mathbf{F}_i - \mathbf{X}\mathbf{c}_i\mathbf{Y}^{\mathrm{T}}\right\|_F^2 \tag{A14}$$

where $\|\cdot\|_{\mathrm{F}}$ represents the Frobenius matrix norm, F_i is the excitation-emission matrix for sample *i* and $\mathbf{c}_i$ is a diagonal matrix containing the corresponding component concentrations. In the single standard addition procedure, *I*=2, $\mathbf{F}_1$=**M** and $\mathbf{F}_2$=**W**, whereas for multiple additions, *I*=*N*+1, $\mathbf{F}_1$=**M** and $\mathbf{F}_{2\ldots N+1}$=$\mathbf{W}^{(1\ldots N)}$. The number of components (a parameter not previously known in complex samples) is estimated as described above for GRAM.

Finally, the SWATLD algorithm alternatively minimizes three objective functions with intrinsic relationship with the trilinear model, and has the advantages of fast convergence and insensitivity to the excess factors used in calculations [15].

In all methods, the standard deviation for the predicted concentration $s(\hat{c}_k)$ was estimated according to the following expression [31]:

$$s(\hat{c}_k) = s_k^{-1}(h_k+1)^{1/2}\sigma \tag{A15}$$

where s_k and h_k are the sensitivity and unknown sample leverage for constituent *k* respectively, and σ denotes the standard deviation of the instrumental noise.

References

1. Brereton RG (2000) Analyst 125:2125
2. Beltrán JL, Ferrer R, Guiteras J (1998) Anal Chim Acta 373:311
3. Wentzell PD, Nair SS, Guy RD (2001) Anal Chem 73:1408
4. Goodman-Hillman A, Rall T, Nier A, Taylor P (1996) The pharmacological basis of therapeutics. MacGraw Hill, New York
5. Borraccetti MD, Damiani PC, Olivieri AC (2002) Anal Chim Acta (in press)
6. Paulus HE, Furst DE, Dromgoole SH (1987) (eds) Drugs for rheumatic disease. Churchill Livingstone, New York
7. Konstantianos DG, Ioannou PC (1996) Analyst 121:909
8. Salinas F, Muñoz de la Peña A, Durán-Merás I, Durán MS (1990) Analyst 115:1007
9. Damiani PC, Ribone ME, Ibañez GA, Olivieri AC (1995) Analyst 120:443
10. Wilson ID, Ismail IM (1986) J Pharm Biomed Anal 4:663
11. Pérez-Ruiz T, Martínez Lozano C, Tomás V, Carpena J (1998) J Pharm Biomed Anal 17:719
12. Sánchez E, Kowaslski BR (1986) Anal Chem 58:496
13. Faber NM, Ferré J, Boqué R (2001) Chemom Intell Lab Syst 55:67
14. Bro R (1997) Chemom Intell Lab Syst 38:149
15. Chen ZP, Wu HL, Jiang JH, Li Y, Yu RQ (2000) Chemom Intell Lab Syst 52:75
16. Muñoz de la Peña A, Ndou TT, Zung JB, Greene KL, Live DH, Warner IM (1991) J Am Chem Soc 113:1572
17. Muñoz de la Peña A, Durán-Merás I, Salinas F, Warner IM, Ndou TT (1991) Anal Chim Acta 255:351
18. Durán-Merás I, Muñoz de la Peña A, Salinas F, Rodríguez Cáceres I (1994) Analyst 119:1215
19. Escandar GM (1999) Analyst 124:587
20. Arancibia JA, Escandar GM (2001) Analyst 126:917
21. Escandar GM (1999) Spectrochim Acta Part A 55:1743
22. Martens H, Naes T (1989) Multivariate calibration. Wiley, Chichester, UK
23. Haaland DM, Thomas EV (1988) Anal Chem 60:1202
24. Goicoechea HC, Olivieri AC (2000) Trends Anal Chem 19:599
25. Faber NM, Ferré J, Boqué R (2001) Chemom Intell Lab Syst 55:67
26. MATLAB 5.3 (1999) The MathWorks Inc., Natick, Massachusetts, USA
27. Bro R (1997) Chemom Intell Lab Syst 38:149
28. González AG, Herrador MA, Asuero AG (1999) Talanta 48: 729
29. Salinas F, Muñoz de la Peña A, Durán MS (1991) Analyst 116: 291
30. Collado MS, Mantovani VE, Goicoechea HC, Olivieri AC (2000) Talanta 52:909
31. Faber NM (2001) J Chemometrics 15:743
32. Faber NM, Kowalski BR (1997) J Chemometrics 11:181