ELSEVIER

Analytica Chimica Acta 409 (2000) 159–170

ANALYTICA CHIMICA ACTA

www.elsevier.com/locate/aca

Application of two- and three-way chemometric methods in the study of acetylsalicylic acid and ascorbic acid mixtures using ultraviolet spectrophotometry

Marcelo M. Sena, Julio Cesar B. Fernandes, Laércio Rover Jr., Ronei J. Poppi, Lauro T. Kubota *

Instituto de Química, UNICAMP, P.O. Box 6154, 13083-970, Campinas, SP, Brazil

Received 23 August 1999; received in revised form 8 December 1999; accepted 27 December 1999

Abstract

In this work, mixtures of acetylsalicylic acid (ASA) and ascorbic acid (AA) were studied by ultraviolet spectrophotometry (210–300 nm) using parallel factor analysis (PARAFAC) and partial least square (PLS). The study was carried out in the pH range from 1.0 to 5.5 and with a concentration range from 1.0×10^{-5} to 1.0×10^{-4} mol l^{-1} of both analytes. PARAFAC was used for spectra deconvolution, pK_a estimation for both acids and to check the presence of salicylic acid (SA), due to the possible ASA decomposition. The estimated first pK_a was equal to 3.41 and 4.10 for ASA and AA, respectively. Multivariate calibration models using PLS at different pH and *N*-way PLS were elaborated for simultaneous determination of ASA and AA in pharmaceutical samples. The best models for the system were obtained with *N*-way PLS2 and PLS2 at pH 1.1. The results obtained for simultaneous determination of ASA and AA in samples were in agreement to the values specified by the manufacturers and the recovery was between 97.6 and 103.6%. Nevertheless, these models failed to predict ASA decomposition to SA in simulated samples. Thus, a new PLS-pH1 model considering SA was built and applied successfully in simulated samples.

Keywords: pK_a estimation; PARAFAC; PLS; Acetylsalicylic acid; Ascorbic acid; Simultaneous determination

1. Introduction

Acetylsalicylic acid (ASA) and ascorbic acid (AA) are widely employed in pharmaceutical formulations and they are probably the major consumed drugs in the world [1,2]. In the last decade, various manufacturers have started to commercialize these substances together, in formulations that combine the action of the ASA for the relief of headaches and fever, with the power of the AA to increase the organism's resistance against microorganisms, as it participates in antibody formation [2]. There are many titrimetric methods for the determination of these substances, such as acid–base titration for ASA [3] or redox titration with iodine, chloramine-T, iodate and dichromate, for AA [3–6]. Other important methods are based on spectrophotometry. For AA a great variety of spectrophotometric methods exists, in which chromophores are employed or not [7–9], while for ASA, the Trinder method (purple complex of salicylate with Fe^{3+}) is well known [10]. HPLC, potentiometry and amperometry for ASA [11–16] and AA [17–23] are

* Corresponding author. Tel.: +55-019-788-3127; fax: +55-019-788-3023.
E-mail address: kubota@iqm.unicamp.br (L.T. Kubota).

0003-2670/00/$ – see front matter
PII: S0003-2670(00)00707-8

also employed. However, at the moment, few methods have been developed for simultaneous determination of ASA and AA. The most recent ones utilize either HPLC [24], a good method, though with high cost, or differential spectrophotometric measurements [25]. Thus, with the increase of the production and consumption of drugs that employ ASA and AA together, it becomes interesting to develop new methods for their simultaneous determination.

On the other hand, due to the sophisticated experimental designs and the increasing amount of data originating from modern instrumentation, the investigation of *N*-dimensional (or *N*-way or *N*-mode) data arrays has attracted much attention. Three-dimensional arrays may be generated by collecting data tables with a fixed set of objects and variables under different experimental conditions, such as sampling time, temperature, pH, etc. The tables obtained under various conditions can be stacked providing a cubic arrangement of data (a parallelepiped whose lines can be objects, whose columns, variables and whose slices, conditions). In many practical relevance situations even higher-dimensional arrays may be considered [26]. The data generated by hyphenated methods, having at least three dimensions, can be considered as one of the most suitable types of data for *N*-dimensional analysis. Among the hyphenated techniques, fluorescence spectroscopy, high-performance liquid chromatography with diode array detection (HPLC-DAD) and gas chromatography with mass spectrometry detection (GC-MS) can be mentioned. Other examples of three-way data can be found in the analysis of kinetic [27], image [28], environmental [29], chemical batch process [30] data.

Several works have appeared in the literature applying PARAFAC to deconvolution problems, such as: the resolution of the spectra of proteins [31], fluoranthenes [32], plant pigment-complexes [33], sugar samples [34], hydrocarbons present in ocean water [35] and isomers of Piroxicam, an anti-inflammatory drug [36]. Ross and Leurgans [37] have presented a good description of some applications of multilinear models in spectroscopy. Finally, it is interesting to point out the da Silva and Novais work [38], who have applied PARAFAC to the decomposition of synchronous fluorescence spectra of mixtures of salicylic, gentisic and salicyluric acid (metabolites of acetylsalicylic acid) as a function of pH and their concentrations.

For multi-way calibration, *N*-way PLS has recently received more attention than other tools. Its applications can be found on HPLC-DAD [39], molecular fluorescence spectrometry [34,40], UV/Visible spectroscopic-kinetic data [41] and 3D-QSAR data [42].

An exploratory study of the simultaneous equilibria of ASA and AA, a verification if ASA could undergo decomposition under the experimental conditions, an estimation of the dissociation constants by the decomposition of the three-way spectral data using PARAFAC, as well as, the development of a multivariate methodology for a rapid simultaneous determination of ASA and AA in pharmaceutical tablets with good precision and accuracy are ascribed.

2. Theory

2.1. Parallel factor analysis (PARAFAC)

PARAFAC, one of several decomposition methods for *N*-way data, is a generalization of principal component analysis (PCA) [43] to higher orders. It can be considered a constrained version of the more general method Tucker3 [26,44,45] with an identity core matrix. It is less flexible, uses fewer degrees of freedom and provides an unique solution independent of rotation. This last feature is a great advantage to the modeling of spectroscopic data. The true underlying spectra (or whatever constitute the variables) will be found if the data is indeed tri-linear, the right number of components is used and the signal-to-noise ratio is appropriated [46]. A PARAFAC model of a three-way array is given by three loading matrices, $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$, with elements a_{if}, b_{jf} and c_{kf} (Eq. (1)), respectively (f=1–F principal components). The tri-linear model is found to minimize the sum of squares of the residues, e_{ijk} in the model [46,47], which is represented as follows:

$$\boldsymbol{X}_{ijk} = \sum a_{if}\, b_{jf}\ \ c_{kf} + e_{ijk} \tag{1}$$

where a_f, b_f and c_f are the fth columns of the loading matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$, respectively.

An important difference between the two-way PCA and the multi-way PARAFAC is that the PARAFAC model is not nested. This fact means that

the parameters of an $F+1$ component model are not equal to the parameters of an F component model plus one additional component. The reason for this is that the components are not required to be orthogonal, hence independent. Therefore, every model has to be calculated specifically with all its components [47]. The algorithm used to solve the PARAFAC model is Alternating least squares (ALS) [26,44,46]. ALS successively assumes the loadings in two modes and then estimates the unknown set of parameters of the last mode. The algorithm converges iteratively until the relative change in fit between two iterations is below a certain value (the default is 1×10^{-6}). It is initialized by either random values or values calculated by a direct tri-linear decomposition based on the generalized eigenvalue problem [46]. Constraining the PARAFAC solution can sometimes be helpful in terms of the interpretability or the stability of the model. The fit of a constrained model will always be lower than the fit of an unconstrained model, but if the constrained one is more interpretable and realistic, this may justify the decrease in fit. The most often used constraints are orthogonality and non-negativity. The resolution of spectra used to require the non-negativity constraint since negative spectral parameters do not make sense [46,47].

2.2. PLS and N-way PLS

The basic principle of the multivariate calibration is the simultaneous utilization of many independent variables, $x_1, x_2, \ldots, x_n$, to quantify one or more dependent variables of interest, y. The partial least squares (PLS) regression analysis is the most widely used method for this purpose, and it is based on the latent variable decomposition relating two blocks of variables, matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ [48], which may contain spectral and concentration data, respectively. These matrices can be simultaneously decomposed into a sum of f latent variables, as follows:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{E} = \sum \boldsymbol{t}_f \boldsymbol{p}_f^t + \boldsymbol{E} \tag{2}$$

$$\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{Q}^{\mathrm{T}} + \boldsymbol{F} = \sum \boldsymbol{u}_f \boldsymbol{q}_f^t + \boldsymbol{F} \tag{3}$$

in which $\boldsymbol{T}$ and $\boldsymbol{U}$ are the score matrices for $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively; $\boldsymbol{P}$ and $\boldsymbol{Q}$ are the loadings matrices for $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively, and $\boldsymbol{E}$ and $\boldsymbol{F}$ are the residual matrices. The two matrices are correlated by the scores $\boldsymbol{T}$ and $\boldsymbol{U}$, for each latent variable, as follows:

$$\boldsymbol{u}_f = b_f \boldsymbol{t}_f \tag{4}$$

in which b_f is the regression coefficient for the f latent variable.The matrix $\boldsymbol{Y}$ can be calculated from $\boldsymbol{u}_f$, as in Eq. (5), and the concentration of the new samples can be estimated from the new scores $\boldsymbol{T}^*$, which are substituted in Eq. (5), leading to Eq. (6)

$$\boldsymbol{Y} = \boldsymbol{T}\boldsymbol{B}\boldsymbol{Q}^{\mathrm{T}} + \boldsymbol{F} \tag{5}$$

$$\boldsymbol{Y}_{\mathrm{new}} = \boldsymbol{T}^*\boldsymbol{B}\boldsymbol{Q}^{\mathrm{T}} \tag{6}$$

In this procedure, it is necessary to find the best number of latent variables, which normally is performed by using cross-validation, based on determination of the minimum prediction error [48]. The difference between PLS1, in which the regression is done for each dependent variable individually ($\boldsymbol{Y}$ is a column matrix), and PLS2, in which all dependent variables are used simultaneously, should be also mentioned.

The N-way PLS is an extension of the PLS regression model to multi-way data. The three-way PLS algorithm decomposes the three-way array $\underline{\boldsymbol{X}}$ into a set of triads. Each triad is equivalent to a latent variable in the bilinear PLS and consists of a score vector, $\boldsymbol{t}$, related to the first mode, and two weight vectors, $\boldsymbol{w}^{\mathrm{J}}$ and $\boldsymbol{w}^{\mathrm{K}}$, related to the other two modes. The model is given by Eq. (7):

$$\boldsymbol{X}_{ijk} = \sum \boldsymbol{t}_{if} \boldsymbol{w}_{jf}^{\mathrm{J}} \boldsymbol{w}_{kf}^{\mathrm{K}} + \boldsymbol{e}_{ijk} \tag{7}$$

where $\boldsymbol{e}_{ijk}$ contains the residues. These vectors are calculated to have the maximum covariance with the unexplained part of the dependent variable, y. The system is modeled by the reduction of the error in the predictions and the result is easy to interpret. A complete description of the algorithm is given by Bro [40].

3. Experimental

3.1. Reagents

ASA, AA, potassium dihydrogenophosphate, hydrochloric acid, potassium chloride and sodium salicylate (SA) were purchased from Merck. All the reagents were of analytical-reagent grade. Two

1.0×10^{-2} mol l^{-1} stock solutions were prepared: AA in water and ASA in ethanol/water 10:90 (ASA has low solubility in water). Five intermediate solutions of each acid were prepared from the stock solutions, in the concentration range from 3.0×10^{-4} to 3.0×10^{-3} mol l^{-1}. Another six solutions were prepared in the pH range from 1.0 to 5.5: four buffer solutions were prepared from potassium dihydrogenophosphate and their pH were adjusted with H_3PO_4/KOH at 3.37, 4.19, 5.00 and 5.51, respectively; two ionic solutions were prepared from potassium chloride and hydrochloric acid, one at pH 1.08 and another that was adjusted with KOH at pH 2.11. All the solutions were prepared in deionized water.

3.2. Instrumentation and software

The pH values were measured on a Corning pH/Ion Analyzer, model 350, calibrated previously with potassium tetraoxalate (1.68), potassium hydrogenphtalate (4.01) and disodium hydrogenphosphate/potassium dihydrogenphosphate (7.00) standards. All measurements were carried out at 25.0°C, in a Quimis thermostated bath, model 214.D2. A Pharmacia Biotech Spectrophotometer, model Ultrospec 2000, equipped with a Peltier device for temperature control, was utilized. A cuvette of 1.00 cm optical path was used for all measurements. The spectrophotometer was coupled to a Pentium 100 MHz (16 Mb RAM) microcomputer and the Swift–Time Drive and Wavescan softwares, versions 1.0, were used for data acquisition.

The data were treated in a Pentium 233 MHz (64 Mb RAM) microcomputer using MATLAB™ software, version 5.2 (The MathWorks, Natick, MA, USA). PARAFAC and *N*-PLS calculus were carried out in 'The *N*-way Toolbox', version 1.00 beta (C. Andersson and R. Bro, Foodtechnology, Kopenhagen, Denmark), while PLS calculus were carried out in the 'PLS Toolbox', version 2.0 (Eigenvector Technologies, Manson, WA, USA).

3.3. Procedures

3.3.1. Experimental design and standard solutions

Fig. 1 shows the experimental design, with two factors (ASA and AA) and five levels (concentration range from 1.0×10^{-5} to 1.0×10^{-4}), used to generate the data array. Seventeen standard solutions were prepared directly inside the cuvette, by the addition of 100 ml of both intermediate acid solutions in 2.80 ml of the respective buffer or ionic solution in each pH. This procedure was repeated for all the pH sets. The spectra of these solutions were scanned in the ultraviolet region, from 210 to 300 nm (step 1 nm). The whole solution set was used for spectral deconvolution and pK_a estimation, with the PARAFAC method. This three-way data array was assembled with 17 solutions in the rows, 91 wavelengths in the columns and 6 pH values in the slices.

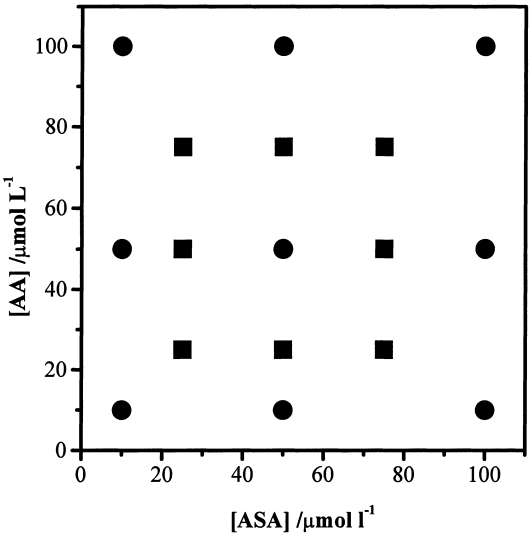

Fig. 1. Experimental design used in the studies of ASA and AA mixtures. (●) Calibration set; (■) Validation set.

For multivariate calibration, the solution set was divided in two sets, one for calibration (9) and another for validation (8), according to Fig. 1. For comparison purposes, the spectra of pure 5.0×10^{-5} mol l^{-1} ASA, AA and SA solutions, at each pH value, were recorded. In this case, only 76 wavelengths were used. to construct PLS models and, therefore, a $9\times76\times6$ data array was used for *N*-PLS.

3.3.2. Assay for kinetic study of the ASA decomposition

In a cuvette, 2.70 ml of a 1.0×10^{-2} mol l^{-1} ASA solution (10% of ethanol/water) and 300 ml of 1.8 mol l^{-1} iron (III) sulfate solution prepared in concentrated nitric acid were added. The kinetics of ASA

hydrolysis to SA was studied, monitoring the formation of a purple complex between SA and Fe^{3+} at 540 nm and 25.0°C, over 30 h (step 5 min), at pH 2–3.

3.3.3. Sample determination and recovery

The pharmaceutical samples from two different manufacturers, containing ASA and Vitamin C together, were acquired in a local drugstore. Each tablet was dissolved in deionized water and diluted to 500.0 ml in a volumetric flask. Then, 50.0 ml of this solution was diluted in a 100.0 ml volumetric flask. Finally, 100 μl of the sample solution was added into a cuvette containing 2.80 ml of the respective buffer or ionic solution with specified pH and 100 μl of deionized water. The recovery method was performed by two additions of 100 μl of a 3.0×10^{-4} mol l^{-1} standard solution of each acid. The spectra were obtained in the same conditions described previously. The analytes concentrations in the samples were determined using the PLS and *N*-PLS models.

3.3.4. Evaluation of ASA decomposition in simulated samples

To verify if the best PLS model is able to predict ASA to SA decomposition, mixture spectra of ASA and AA (both 5.0×10^{-5} mol l^{-1}) under 10, 20, 30, 40, 50 and 100% of decomposition were simulated through linear combination from the pure spectra experimentally obtained at pH 1. As this model failed to predict the decomposition, a new PLS-pH1 calibration model was built, aiming at modeling also SA. A new set of mixtures with an ASA decomposition degree between 5 and 35% were simulated for this purpose, according to Table 1. These spectra were used together with the original 17 real solutions without decomposition, totalizing 33 solutions (nine original ones) for calibration and 27 solutions for validation (eight original ones) of the model. This new model was used to predict the simulated set of samples under ASA decomposition (10–100%) and the pharmaceutical samples.

4. Results and discussion

4.1. Exploratory study of the system using PARAFAC

Fig. 2 shows the experimentally obtained spectral surfaces. The most remarkable feature in this figure is the shift of the peak at about 240 nm to higher wavelengths, as the pH increases. Fig. 3 shows the spectra of the pure solutions of AA, ASA and SA (5.0×10^{-5} mol l^{-1}) at different pH. As it can be observed, AA presents λ_{max} at 243 nm in acid solutions (pH<2) and at 265 nm in solutions at pH higher than 5. On the other hand, the observed spectra of ASA solutions present λ_{max} at 229 nm at pH<2 and between 220 and 230 nm at pH higher than 4. However, ASA may undergo hydrolysis, yielding SA and/or salicylate [49], hence causing a spectrum shift. This could explain the deformation of the ASA spectrum at pH higher than 4, showing no clear maximum of absorbance. Then, PARAFAC was utilized to study the behavior of the ionic equilibrium of both acids and to verify if ASA could suffer decomposition under the experimental conditions, through spectra deconvolution and pK_a estimation.

The data was arranged in a three-way array $17\times91\times6$, composed of 17 solutions, with different AA and ASA concentrations, in the rows, 91 wavelengths in the columns and 6 pH values in the slices. No preprocessing (centering or auto scaling) was applied to the data. The non-negativity constraint was imposed to the three modes, since neither absorbance nor concentrations is expected to be negative. Random orthogonalized values were used for PARAFAC initialization and the convergence criterion was 1×10^{-6}. In addition to the data set, PARAFAC requires the definition of the numbers of factors to be included in the model. Among the ways of determining this correct number [46], intuition and comparison of the results with the external knowledge of the data being modeled were used. Hence, if the ASA decomposition to SA has not occurred, a four-component PARAFAC model should give the correct solution.

It was tried to run PARAFAC with several number of components. Five and six-component models provided degenerated solutions and a three-component model left systematic variance in the residuals, thus indicating that more components can be extracted. A four-component PARAFAC decomposition yielded a stable solution, that was considered the best model. Fig. 4a shows the loadings of the second mode of this model. The obtained curves match exactly the expected spectra (see Fig. 3) for AA, ASA and their monodissociated species. Fig. 4b shows the loadings of the third mode, which provided the species distri-

Table 1
Simulated samples used in the construction of the PLS model for simultaneous determination of the AA, ASA and SA[a]

Number	Concentration		
	10^5 [AA] (mol l^{-1})	10^5 [ASA] (mol l^{-1})	10^5 [SA] (mol l^{-1})
#1[b]	1.00	0.95	0.05
#2	5.00	0.95	0.05
#3	10.00	0.95	0.05
#4	1.00	4.75	0.25
#5	5.00	4.75	0.25
#6	10.00	4.75	0.25
#7	1.00	9.50	0.50
#8	5.00	9.50	0.50
#9	10.00	9.50	0.50
#10	1.00	0.85	0.15
#11	5.00	0.85	0.15
#12	10.00	0.85	0.15
#13	1.00	4.25	0.75
#14	5.00	4.25	0.75
#15	10.00	4.25	0.75
#16	1.00	8.50	1.50
#17	5.00	8.50	1.50
#18	10.00	8.50	1.50
#19	1.00	0.65	0.35
#20	5.00	0.65	0.35
#21	10.00	0.65	0.35
#22	1.00	3.25	1.75
#23	5.00	3.25	1.75
#24	10.00	3.25	1.75
#25	1.00	6.50	3.50
#26	5.00	6.50	3.50
#27	10.00	6.50	3.50

[a] Decomposition levels: 5%, 15% and 35%.
[b] Odd numbers: calibration and even numbers: validation.

bution as a function of the solution pH. It was possible to estimate the dissociation constants for both acids through the crossing lines in this figure. The first pK_a was estimated as equal to 3.41 and 4.10 for ASA and AA, respectively. The estimatives were considered good in comparison to the literature values (3.49 — ASA and 4.17 — AA, at 25°C) [49], which are determined by potentiometric titration [50]. The relative errors were 2.3% for ASA and 1.7% for AA. This four-component model demonstrated that ASA was not decomposed.

A kinetic study of the ASA hydrolysis was carried out to confirm this assumption. Fig. 5 shows the decrease in the ASA concentration (1×10^{-2} mol l^{-1}) monitored over a period of 30 h, jointly with the linearization plot for pseudo first order kinetics. The estimated kinetic constant was 7.4×10^{-6} s^{-1} and the half-life was 26 h. This result also indicates that ASA was not significantly decomposed during the spectra acquisition. Moreover, it is important to emphasize that the more diluted and more acid the solutions, the slower the decomposition.

4.2. Multivariate calibration and simultaneous determination using PLS and N-PLS

The multivariate calibration is a powerful tool for simultaneous determinations, because it extracts more information from the data and allows to build more robust models. So, it makes possible analysis without the use of separation or extraction methods.

A PARAFAC model can be used for a multivariate calibration of three-way data following the same strategy of PCR [46]. The loadings of the first PARAFAC

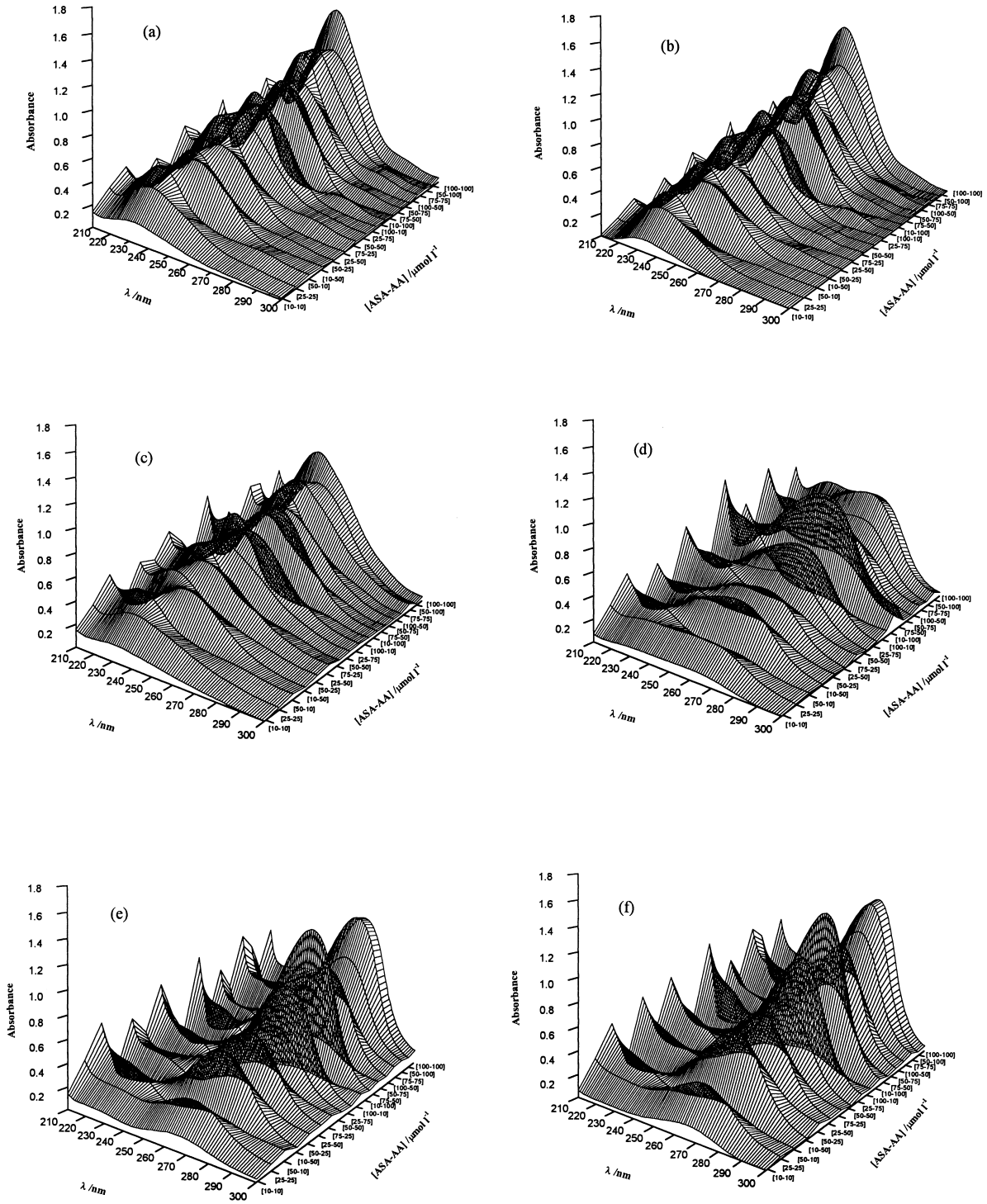

Fig. 2. Spectral surfaces obtained for the data set used in the study of ASA and AA mixtures by multivariate methods. pH: (a) 1.08; (b) 2.11; (c) 3.37; (d) 4.19; (e) 5.00; (f) 5.51.

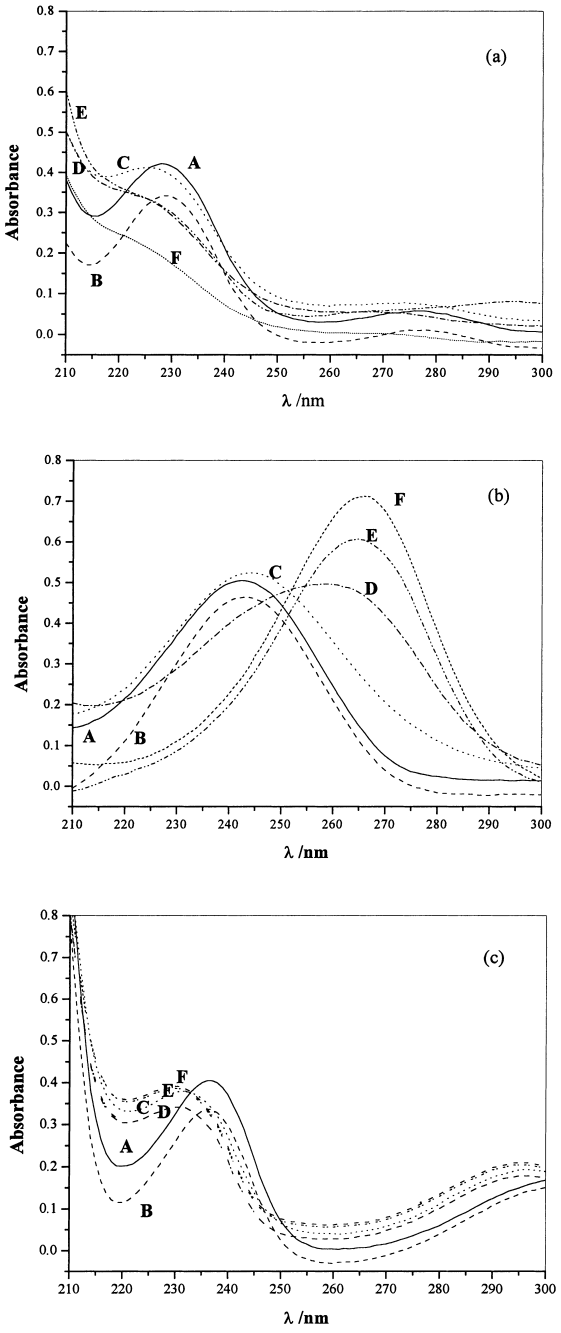

Fig. 3. Pure spectra obtained for (a) ASA, (b) AA and (c) SA solutions. Concentration: 5.0×10^{-5} mol l^{-1}. pH: A. 1.08; B. 2.11; C. 3.37; D. 4.19; **E**. 5.00; **F**. 5.51.

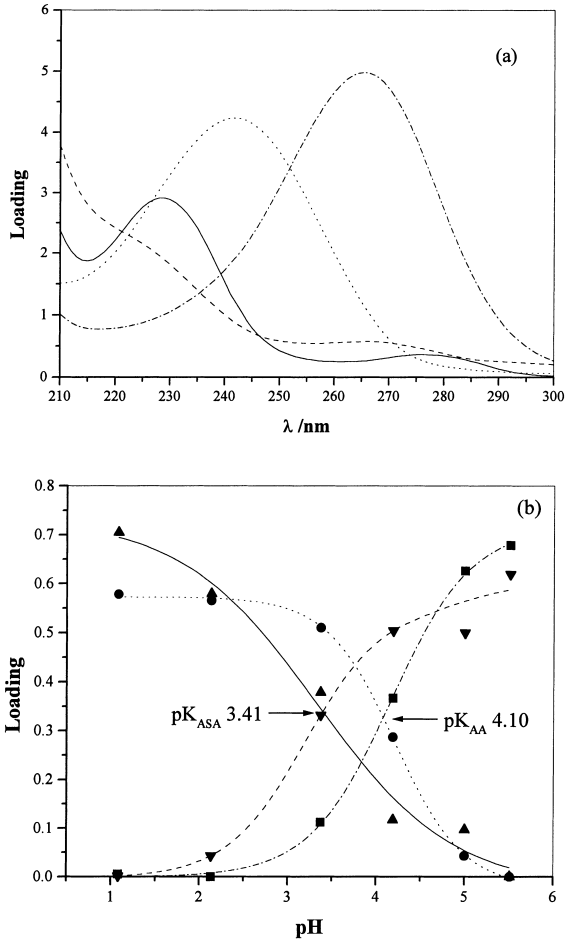

Fig. 4. (a) Deconvoluted spectra and (b) distribution of the species (ASA–AA) as a function of the pH, obtained from the PARAFAC method. (—) ASA, (- -) ASA monoanion, (· · ·) AA and (-●-) AA monoanion.

mode can be used as independent variables in a multiple linear regression. However, the use of *N*-PLS for this purpose is more advantageous and used to provide better results [40,47]. The use of PARAFAC is too laborious because it is necessary to build and test a separate model for each number of components, since the PARAFAC solution is not nested. Besides, the PARAFAC algorithm can be very slow, depending on the number of variables and *N*-PLS might stabilize the predictive model by incorporating the dependent variables in the decomposition of independent variables.

Therefore, it was decided to perform a multivariate calibration using *N*-PLS and compare it with

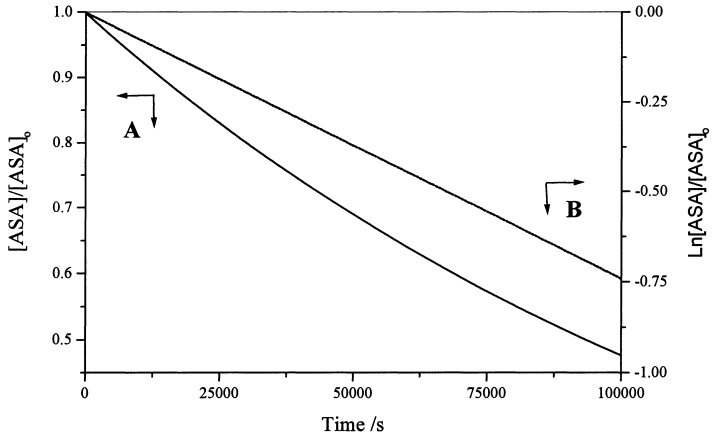

Fig. 5. A. Kinetics curve of the ASA hydrolysis. B. Linearization plot used to estimate the kinetic constant and the half-life.

two-way PLS models built for each pH value individually. According to an experimental design (Fig. 1), nine solutions were used to construct the models (calibration set) and another eight solutions to validate them (validation set). The models were validated using cross-validation. The root mean square error of prediction (RMSEP) and the correlation coefficient (r) between the real and the predicted values were used as parameters for comparison among the models. PLS1 and PLS2 models were performed and no significant differences were observed between the results. Then, PLS2 was chosen because it is simpler to perform.

The spectral region from 210 to 225 nm was eliminated, improving the sample predictions. The buffers absorb strongly in this region, introducing errors in the analysis due to background fluctuations. Hence, the elimination of this region turns the models more robust. Nevertheless, this region was useful, providing essential information for PARAFAC modeling, since it was not possible to recover the correct spectra without it. Therefore, 9×76 matrices were used to construct PLS models and a $9\times76\times6$ data array was used for N-PLS.

Table 2 presents the RMSEP and r values for all calculated models. All the PLS models were built using two latent variables, while for N-PLS four latent variables were used. The best calibration models were obtained using N-PLS and PLS at pH lower than 2. Curiously, this would not be expected, because the spectra of ASA and AA present better resolution and sensitivity ($\varepsilon_{265}>\varepsilon_{243}$ for AA) at pH>3 (Fig. 3). However, the observed behavior can be explained by the instability of the monoascorbate species. It was shown that about 11–21% of AA undergoes decomposition to dehydroascorbic acid (does not absorb at 210–300 nm) in the presence of oxygen at pH 5.0–5.6, in a period of 30–60 min [51]. Besides, the ASA spectra are deformed in pH>3, in that acetylsalicylate does not present a defined maximum in a given wavelength

Table 2
Errors of prediction and regression coefficients between the real and the predicted values obtained for the validation set, for each proposed multivariate calibration model

RMSEP and r	$PLS2_{pH}$						N-way PLS2
	1.08	2.11	3.37	4.19	5.00	5.51	
10^6 [ASA_{RMSEP}] (mol l^{-1})	8.21	4.55	10.3	10.7	19.2	26.5	3.77
ASA_r	0.991	0.998	0.987	0.996	0.965	0.953	0.998
10^6 [AA_{RMSEP}] (mol l^{-1})	5.17	7.13	10.8	10.0	25.6	7.42	2.42
Aa_r	0.998	0.997	0.991	0.995	0.923	0.997	0.999

Table 3
Results obtained from simultaneous determination of ASA and AA in tablets, using multivariate calibration

Sample[a]	Predict values (g per tablet)			
	N-way PLS2		PLS2-pH1	
	ASA	AA	ASA	AA
#1	0.392 (2.3%)[b]	0.245 (2.4%)	0.416 (2.2%)	0.254 (2.4%)
#2	0.397 (2.3%)	0.252 (2.3%)	0.396 (2.3%)	0.248 (2.4%)

[a] Tablet composition: #1 contains 0.40 g ASA, 0.24 g AA, sodium bicarbonate, starch and lemon flavor additive. #2 contains 0.40 g ASA, 0.24 g AA, sodium bicarbonate, starch, red dye and orange flavor additive.
[b] Relative standard deviation of three determinations.

and, thus, small shifts in the absorbance might provoke large errors in the predicted concentrations.

The two best calibration models, N-PLS and PLS at pH1, were utilized for simultaneous determination of ASA and AA in two different tablet samples. The results are listed in Table 3 and are in agreement with those specified by the manufacturers. The recovery method for two additions of ASA and AA were also performed, corroborating the efficiency of the multivariate calibration. Table 4 presents the percentage of recovery for each addition.

Although *N*-PLS has shown the best model, PLS model at pH1 can be considered a more suitable model for a routine analysis, because *N*-PLS requires the samples are measured at several pH. Furthermore, the predictions and the recoveries from *N*-PLS were only slightly better than the PLS ones at pH1.

4.3. Construction of a PLS-pH1 model for estimating ASA decomposition in simulated samples

ASA determination used to be performed indirectly, through the conversion of ASA to SA, in a back acid–base titration or in a further spectrophotometric determination of the Fe^{3+}–SA complex [3,10]. However, these methods can not provide a good quantification of ASA decomposition in the samples, that can be provoked mainly by the moistness of the environment. If the decomposition occurs, the drugs can be commercialized in lower quantity than the specified and cause problems to the organism due to salicylate toxicity [1]. In addition, it has been shown that the analgesic action mechanism of the aspirin-like drugs is due to acetyl group, which inhibits the cycloxygenase enzyme [52]. Thus, it is interesting to develop methods that take into account the possible decomposition in ASA determination, i.e., that are able to truly quantify ASA.

Table 4
Recovery values obtained for two standard additions of ASA and AA in the pharmaceutical samples

Samples	*N*-way PLS2				PLS2-pH1			
	Added quantity							
	0.03 μmol		0.06 μmol		0.03 μmol		0.06 μmol	
	Recovery (%)							
	ASA	AA	ASA	AA	ASA	AA	ASA	AA
#1	99.9	98.0	99.8	97.6	100.6	101.7	100.5	103.6
#2	100.1	99.9	99.9	99.4	99.7	98.7	98.8	98.1

The PLS-pH1 model obtained in the preceding section was tested in simulated samples under 10–100% of ASA decomposition and failed to predict them. It predicted 2.5% for a solution under 10% and about 30% for a solution under 100% of decomposition. As the decomposition increases, the error in AA prediction increases (20% for total decomposition). So, this model was not able to distinguish between ASA and SA, which can be understood by observing the similarity between the respective spectra in Fig. 3.

In situations where the determination of an analyte is made difficult by the presence of one or several other constituents, instead of eliminating the interfering species, e.g. by a separation procedure, these interferences should be possible to quantify along the primary analyte [41]. Therefore, a new PLS-pH1 model was built to model SA. To construct this new model, 43 new samples were simulated, according to Table 1. 24 simulated plus nine real solutions were used as the calibration set and 19 simulated plus eight real solutions were used as the validation set. The parameters

Table 5
Errors of prediction and regression coefficients obtained with the PLS2-pH1 model that model also SA, for the new validation set[a]

	ASA	AA	SA
10^6 [RMSEP] (mol l^{-1})	6.23	4.89	3.49
r	0.999	0.999	0.997
Sample determination (g per tablet)			
#1	0.426 (2.2%)[b]	0.258 (2.4%)	n.d.[c]
#2	0.401 (2.3%)	0.251 (2.4%)	n.d.[c]

[a] It was also included the sample determinations with this model.
[b] Relative standard deviation of three determinations.
[c] Non-detectable.

Table 6
Samples prediction for the simulated ASA decomposition set

Decomposition degree	10^{-5} Concentration (mol l^{-1})					
	Real			Obtained		
	AA	ASA	SA	AA	ASA	SA
10%	5.00	4.50	0.50	5.00	4.53	0.46
20%	5.00	4.00	1.00	5.00	4.03	0.97
30%	5.00	3.50	1.50	4.99	3.52	1.49
40%	5.00	3.00	2.00	4.99	3.01	2.01
50%	5.00	2.50	2.50	4.98	2.51	2.51
100%	5.00	0.00	5.00	4.96	n.d.[a]	5.09

[a] Non-detectable.

of this model (RMSEP and *r*) are shown together with the results for determination in tablet samples, in Table 5. The model was considered very good and was able to predict ASA to SA decomposition with large accuracy, as shown in Table 6.

5. Conclusions

The spectral deconvolution using traditional two-way chemometric methods, such as PCA, used to be impossible unless hard constraints were imposed on the model. Due to its unique solution, the three-way PARAFAC model demonstrated to be an excellent tool to study ASA and AA equilibria, allowing spectral deconvolution and confirming the absence of ASA decomposition under the investigated conditions. Furthermore, PARAFAC and UV spectrophotometry provided a new way to estimate p*K* values, that can be applied to very weak acid or alkaline substances. In attempt to achieve more accurate estimations, the procedure can be repeated taking more pH measurements in the regions closer to the expected pK_a values.

On the other hand, few methods for simultaneous determination of ASA and AA have been developed until the present. Besides, these methods are too laborious or too expensive, requiring for example, chromatographic separation [24]. *N*-way PLS or PLS multivariate calibration using UV spectrophotometric data can be considered a suitable method for a rapid determination. Although multi-way models have presented better results than bilinear models, PLS at pH1 was considered simpler and good enough to be applied for routine analysis. Alternatively, a multiple linear regression with variable selection could also be tried for simultaneous determination of ASA and AA in pharmaceutical samples.

Finally, the model can be used directly for ASA determination without need of SA conversion. The modeling of SA allowed to obtain a method robust enough to determinate ASA in the presence of SA.

Acknowledgements

The authors thank Dr. Rasmus Bro and Dr. Clauss Andersson for ‘The *N*-way Toolbox’, Professor Carol H. Collins for English revision and FAPESP and CNPq for the financial support.

References

[1] J.J. Thiessen, Acetylsalicylic Acid: New Uses for an Old Drug, in: H.J.M. Barnett, J. Hirsh, J.F. Mustard (Eds.), Raven Press, New York, 1992, pp. 49–61.
[2] C.A.B. Clemetson, Vitamin C, CRC Press, Boca Raton, FL, 1989.
[3] British Pharmacopoeia, Her Majesty’s Stationary Office, London, 1980.
[4] M.A. Koupparis, T.P. Hadjiioannou, Anal. Chim. Acta 94 (1977) 367.
[5] A.C. Calokerinos, T.P. Hadjiioannou, Microchem. J. 28 (1983) 464.
[6] P.C. Gritzapis, C.E. Efstathiou, T.P. Hadjiioannou, Anal. Chim. Acta 171 (1985) 165.
[7] Y. Fung, S. Luk, Analyst 110 (1985) 1439.
[8] K.K. Verna, A. Jain, A. Verna, A. Chaurasia, Analyst 116 (1991) 641.
[9] K.K. Verna, A. Jain, A. Chaurasia, Talanta 42 (1995) 779.

[10] P. Trinder, Biochem. J. 57 (1954) 301.
[11] A.M. DiPietra, R. Gatti, V. Andrisano, V. Cavrini, J. Chromatogr. A 729 (1996) 355.
[12] G.P. McMahon, M.T. Kelly, Anal. Chem. 70 (1998) 409.
[13] L.T. Kubota, J.C.B. Fernandes, L. Rover Jr., G.O. Neto, Talanta 50 (1999) 661.
[14] M. Neumayr, O. Friedrich, G. Sontag, F. Pittner, Anal. Chim. Acta 273 (1993) 469.
[15] M. Ehrendorfer, G. Sontag, F. Pittner, Fresenius' J. Anal. Chem. 356 (1996) 75.
[16] J.M.P.J. Garrido, J.L.F.C. Lima, C. Delerue-Matos, V.V.M. Meijden, Port. Electrochim. Acta 15 (1995) 335.
[17] L.W. Doner, K.B. Hicks, Anal. Biochem. 115 (1985) 225.
[18] P. Wimalasiri, R.B.H. Wills, J. Chromatogr. 256 (1983) 368.
[19] P. Janda, J. Weber, L. Dunsch, A.B.P. Lever, Anal. Chem. 68 (1996) 960.
[20] J.C.B. Fernandes, L.T. Kubota, G.O. Neto, Anal. Chim. Acta 385 (1999) 3.
[21] J.C.B. Fernandes, L.T. Kubota, G.O. Neto, Electroanalysis 11 (1999) 475.
[22] G.M. Greenway, P. Ongomo, Analyst 115 (1990) 1297.
[23] S. Ikeda, H. Satake, Y. Kohri, Chem. Lett. 6 (1984) 873.
[24] V. Kmetec, J. Pharm. Biom. Anal. 10 (1992) 1073.
[25] H.N. Dogan, A. Duran, Pharmazie 53 (1998) 781.
[26] R. Henrion, Chemom. Intell. Lab. Syst. 25 (1994) 1.
[27] P. Hindmarch, K. Kavianpour, R.G. Brereton, Analyst 122 (1997) 871.
[28] P. Geladi, Chemom. Intell. Lab. Syst. 5 (1989) 209.
[29] R. Henrion, G. Henrion, G.C. Onuoha, Chemom. Intell. Lab. Syst. 16 (1992) 87.
[30] K.S. Dahl, M.J. Piovoso, K.A. Kosanovich, Chemom. Intell. Lab. Syst. 46 (1999) 161.
[31] S.R. Durrel, C.W. Lee, R.T. Ross, E.L. Gross, Arch. Biochem. Biophys. 278 (1990) 148.
[32] D.S. Burdick, X.M. Tu, L.B. McGown, D.W. Millican, J. Chemometr. 4 (1990) 15.
[33] R.T. Ross, C.H. Lee, C.M. Davis, B.M. Ezzeddine, E.A. Fayyad, S.E. Leurgans, Biochim. Biophys. Acta 1056 (1991) 317.
[34] R. Bro, Chemom. Intell. Lab. Syst. 46 (1999) 133.
[35] K.S. Booksh, A.R. Muroski, M.R. Myrick, Anal. Chem. 68 (1996) 3539.
[36] J.A. Martins, M.M. Sena, R.J. Poppi, F.B.T. Pessine, Appl. Spectrosc. 53 (1999) 510.
[37] R.T. Ross, S.E. Leurgans, Stat. Sci. 7 (1992) 289.
[38] J.C.G.E. da Silva, S.A.G. Novais, Analyst 123 (1998) 2067.
[39] K.D. Zissis, R.G. Brereton, S. Dunkerley, R.E.A. Escott, Anal. Chim. Acta 384 (1999) 71.
[40] R. Bro, J. Chemometr. 10 (1996) 47.
[41] A.K. Pettersson, B. Karlberg, Anal. Chim. Acta 354 (1997) 241.
[42] J. Nilsson, S. de Jong, A.K. Smilde, J. Chemometr. 11 (1997) 511.
[43] K.R. Beebe, B.R. Kowalski, Anal. Chem. 59 (1987) 1007A.
[44] P.M. Kroonenberg Principal Component Analysis: Theory and Applications, DSWO, Leiden, 1983.
[45] A.K. Smilde, Chemom. Intell. Lab. Syst. 15 (1992) 143.
[46] R. Bro, Chemom. Intell. Lab. Syst. 38 (1997) 149.
[47] R. Bro, The *N*-way On-line Course on PARAFAC and PLS, http://newton.mli.kvl.dk/htmlcourses/parafac, 1998.
[48] H. Martens, T. Naes, Multivariate Calibration, Wiley, Chichester, 1989.
[49] The Merck Index: An Encyclopedia of Chemicals, Drugs and Biologicals, 12th Edition, Merck & Co., Whitehouse Station, 1996.
[50] M.L.S.S. Gonçalves, Métodos Instrumentais para Análise de Soluções — Análise Quantitativa, Fundação Calouste Gulbenkian, Lisboa, 1983.
[51] J.C.B. Fernandes, G.O. Neto, L.T. Kubota, Anal. Chim. Acta 366 (1998) 11.
[52] I. Wells, L.J. Marnett, Biochemistry 31 (1992) 9520.