Multivariate curve resolution of pH gradient flow injection mixture analysis with correction of the Schlieren effect

Renato L. Carneiro,[a] Jez Willian B. Braga,[a] Ronei J. Poppi[a] and Romà Tauler*[b]

Received 13th December 2007, Accepted 19th February 2008
First published as an Advance Article on the web 13th March 2008
DOI: 10.1039/b719245b

Multivariate curve resolution using alternating least squares (MCR-ALS) was used to quantify ascorbic (AA) and acetylsalicylic (ASA) acids in four pharmaceutical samples using a flow injection analysis (FIA) system with pH gradient and a diode array (DAD) spectrometer as a detector. Four different pharmaceutical drugs were analyzed, giving a data array of dimensions 51 × 291 × 61, corresponding respectively to number of samples, FIA times and spectral wavelengths. MCR-ALS was applied to these large data sets using different constraints to have optimal resolution and optimal quantitative estimations of the two analytes (AA and ASA). Since both analytes give an acid–basic pair of species contributing to the UV recorded signal, at least four components sholuld be proposed to model AA and ASA in synthetic mixture samples. Moreover, one additional component was needed to resolve accurately the Schlieren effect and another additional component was also needed to model the presence of possible interferences (like caffeine) in the commercial drugs tablets, giving therefore a total number of 6 independent components needed. The best quantification relative errors were around 2% compared to the reference values obtained by HPLC and by the oxidation–reduction titrimetric method, for ASA and AA respectively. In this work, the application of MCR-ALS allowed for the first time the full resolution of the FIA diffusion profile due to the Schlieren effect as an independent signal contribution, suggesting that the proposed MCR-ALS method allows for its accurate correction in FIA-DAD systems.

1. Introduction

The application of flow injection analysis (FIA) to the development of analytical methods has became a common procedure due their known advantages of simplicity, feasibility, reproducibility, low reagent consumption, high degree of automation and sampling frequency.[1] When the signal of a FIA system is monitored by a spectrophotometer with a multiwavelength diode array detection (DAD) or an infrared equipment with a Fourier transform (FT-IR) configuration, a large amount of data ordered in a data table or matrix is acquired in a simple and fast way, as spectra can be obtained as a function of time.

The data acquired in these FIA systems can be used in the study of several chemical systems. However, when a mixture of compounds are present in the sample, interferences between compounds can make the direct study or analysis difficult or impossible. In addition, the occurrence of the Schlieren effect, which is a consequence of the light refraction in regions where refractive index gradients are present,[2] and that it has been already described in FIA systems by Krug *et al.*,[3] constitutes a source of variation that may influence the signal-to-noise ratio and the measurement repeatability, therefore producing an increase of errors in the estimates of the concentration of the analyte.

Whenever a data matrix is measured for each individual sample by FIA, second-order calibration methods can be applied, which allow for the determination of the concentration of the analyte in a mixture or complex sample, even in the presence of unknown interferences. This is known in the chemometrics literature as the second-order advantage.[4] The multivariate curve resolution-alternating least squares method (MCR-ALS) can be used as a second-order calibration method, since it can provide both qualitative and quantitative information about the analytes in the sample, in the presence of unknown interferents. Moreover MCR-ALS works adequately with analytes presenting several equilibrium species where deviations of the expected bilinear structure of the data do not occur if the proper number of components selected for the analysis equals the real number of species at equilibrium.[5–10] Their successful application to FIA systems was already proved in previous papers, such as in the determination of mixtures of diprotic organic acids by FT-IR,[5,6] in the quantitative determination of mixtures of amino acids,[7,8] of nucleic acids[9,10] and in the estimation of the pK_a values[11] of a mixture of nucleic acids by UV absorption spectrophotometry.

In this article, the application of MCR-ALS for the simultaneous determination of acetylsalicylic (ASA) and ascorbic acids (AA) in pharmaceutical tablet samples by FIA with pH gradient and DAD in the UV region is described. Moreover the

[a]*Universidade Estadual de Campinas, Instituto de Química, C.P. 6154, CEP, 13084-971, Campinas, SP, Brazil*
[b]*Department of Environmental Chemistry, Institute of Chemistry and Environmental Research–CSIC Jordi Girona, 16, 08034, Barcelona, Spain. E-mail: rtaqam@iiqab.csic.es*

influence of the Schlieren effect is investigated and resolved in the determination of these two analytes, and special attention is paid to the constraints necessary to achieve an optimal resolution among the acid and basic forms of the two analytes, the Schlieren effect and the sample interferences. Four different pharmaceutical drugs were analyzed, and their quantification for the two analytes, ASA and AA, was compared with appropriated reference methods using HPLC and an oxidation–reduction titrimetric method.

2. Experimental

2.1. Reagents

Calibrated volumetric flashes and ultrapure water (Milli-Q, Millipore) were used for all solutions. Acid carrier solution (H_3PO_4, 0.01 mol L^{-1}) was prepared from 85% (w/w) phosphoric acid solution (Synth) and the basic solution (Na_2HPO_4, 0.05 mol L^{-1}) from a 99.0% sodium hydrogen phosphate (Fluka). Standard solutions for calibration and validation were prepared from analytical grade ascorbic acid (AA) 99.75% (Fluka), acetylsalicylic acid (ASA) 99.9% (Synth) and caffeine 99.9% (Synth).

2.2. Flow injection analysis system

Data acquisition was accomplished with a FIA system that was composed of three solenoid valves (N-Research), a power source, a peristaltic pump (Ismatec IPC), tygon tubes of various inner diameters, polytetrafluoroetilene tubes of 0.80 mm inner diameter, T-junctions and a mixture chamber of 0.950 mL. The output of the system was connected to a flow cell with 10 mm of pathlength placed in a Hewlett-Packard HP8452 spectrophotometer equipped with a diode array detector (DAD) with a resolution of 2 nm and adjusted to acquire one spectra per second.

Samples were pumped continuously, and after mixing with water and H_3PO_4 acid carrier, the concentration was approximately 11 times lower than in the original sample, with a flow rate of 0.90 mL min^{-1}. A pH gradient covering approximately from pH 2.0 to pH 7.5 was obtained by injection of 160 μL of a K_2HPO_4 basic solution into the system.

2.3. Synthethic mixtures for calibration and validation

Calibration and validation synthetic mixtures were prepared daily from 400.0 mg L^{-1} solution of ASA, 240.0 mg L^{-1} of AA and 50.0 mg L^{-1} of caffeine (used as interferent only in synthetic validation mixtures). Due to the high decomposition presented by AA, the solutions (of standards and samples) were prepared in Ultrapure water saturated with nitrogen gas and maintained at approximately 10 °C until their analysis, and measured immediately after preparing.

The calibration samples were composed of 11 sample solutions, formed by 9 synthetic mixture samples of ASA and AA, following a composite central design, and by 2 additional samples containing just one of the two analytes (although each of these two analytes gives two acid–base species, both of them are spectrophotometrically active, see later). Three independent replicates of these 11 samples were analyzed, therefore giving a total number of 33 samples for the calibration. The total concentration ranges of the two analytes were established based on their absorption in the ultraviolet region and varied between 0 and 136.4 mg L^{-1} for ASA and between 0 and 82.0 mg L^{-1} for AA.

The validation samples were composed of 5 synthetic mixtures of ASA and AA, in which caffeine was added as interference at a constant concentration of 5.00 mg L^{-1}. This concentration level of caffeine was chosen based on its usual concentration present in the commercial pharmaceutical drug named Doril®. Three replicates of them were also analyzed, giving a total number of 15 validation samples.

2.4. Pharmaceutical samples

Pharmaceutical products analyzed were: Aspirina® +C (Bayer S.A.), Melhoral® C (DM Indústria Farmacêutica LTDA), Doril® (DM Indústria Farmacêutica LTDA) and Sandoz® (Novartis Biociências S.A.). The first two pharmaceutical products contained the two analytes as active compounds and other possibly excipient interferences. Doril® contains ASA and caffeine as active compounds, while Sandoz® presents just AA as active compound and possible excipient interferences.

For the preparation of these pharmaceutical samples, 10 units of each of the four pharmaceutical products were mixed, weighted and grounded. Then, a specific weight of these pharmaceutical product mixtures was dissolved in 500.0 mL volumetric flash and 50.00 mL of this solution were diluted up to 100.00 mL in a volumetric flash. Additionally, with the aim to perform a recovery study, ASA and/or AA were added at three distinct levels to each one of the previously prepared pharmaceutical sample solutions that contained ASA and/or AA. In total, six different samples were prepared for each pharmaceutical product (3 without and 3 with the addition of these two analytes) and all these samples were analyzed in triplicate in the FIA system, producing 18 matrices for each one of the four pharmaceutical products investigated in this work and a total number of 72 (18 × 4) pharmaceutical samples were analyzed by FIA.

2.5. Reference methods

Reference methods used for validation of the MCR-ALS determinations of ASA and AA by FIA were performed by HPLC and iodimetric titration, respectively. The sample preparation was the same as described in the previous section and the determinations were performed in triplicate (for each pharmaceutical product).

For ASA determination, a Shimadzo Prominence high performance liquid chromatography equipped with an APD-M20A diode array detector, a SL20A autosampler and a Microsorb MV C18 5 μ column (250 mm × 4.6 mm) from Varian were used. Separations were carried out with 15 : 85 (v/v) acetonitrile : water as mobile phase, the water was acidified to pH 3.0 with phosphoric acid and a flow rate of 1.0 mL min^{-1} was used. Analytical curves were established with six standards and the chromatograms acquired at 232 nm.

For AA determination, the standard method suggested by the United States Pharmacopoeia[12] was used, which consists of the

titration of AA in an acidified solution with an iodine standard solution and starch as the indicator.

3. Method

3.1. Data matrix arrangements

Every sample analyzed by the FIA-DAD gives a data matrix, **D**, in which the successive DAD spectra acquired at different times during the FIA process are arranged in the rows of this data matrix. The columns of this data matrix describe the time evolution of the absorbance measurements at the different wavelengths. Therefore, the dimensions of each matrix are 291 times and 61 wavelengths, **D**(291,61). In Fig. 1a and 1b, two different data matrix arrangements used in this work for the simultaneous analysis of the different type of samples is summarized. These new data matrix arrangements are called column-wise data matrix augmentation. These new augmented data matrices ($\mathbf{D}_{cal}$ and $[\mathbf{D}_{cal};\mathbf{D}_{pred}]$ in Fig. 1a and 1b, respectively) will have the same number of columns (J wavelengths) than the single sample data matrices, but they will have a larger number of rows, equal to the number of individual data matrices included in the simultaneous analysis (*e.g.* 33 matrices for calibration set) times the number of FIA acquisition times considered in the analysis of each individual sample or FIA run. Therefore the dimensions of these new augmented data matrices will be **D**(33 × 291,61) for calibration set, shown in Fig. 1a.

3.2. Multivariate curve resolution-alternating least squares (MCR-ALS) method

MCR-ALS is used in this paper as a second-order resolution and calibration method, since a data matrix per sample is analyzed to provide both qualitative and quantitative information about the analytes under study in the presence of uncalibrated interferences.[5–11,13] The main goal of the MCR-ALS method is the mathematical resolution of the mixture of signals present in the experimental data matrix **D** into the pure contributions of all significant components and species, according to a bilinear model expressed by the equation:[13,14]

$$\mathbf{D} = \mathbf{C}\mathbf{S}^{T} + \mathbf{E} \quad (1)$$

where **D** (I, J) is the experimental data matrix that contains absorbance measurements at J (61) wavelengths during the I (291) discrete times that the flow analysis process lasts, **C** is the (I, NC) matrix with the concentration profiles of the NC significant components, contributions or species, $\mathbf{S}^{T}$ is the (J, NC) transpose matrix of the pure spectra profiles of these components, and **E** is the (I, J) matrix of the residuals not explained by the modeled components. For the successful application of MCR-ALS, the following are required: (1) an estimation of the number of significant compounds (NC); (2) an initial estimation of either the concentration or spectral profiles of these components; (3) the selection of appropriate constraints for the ALS minimization; (4) the ALS optimization itself; and (5) the extraction and interpretation of the qualitative and quantitative information present from the resolved profiles in **C** and **S** matrices.[5,6,14] MCR-ALS has been discussed in previous references,[5–14] thus only a brief description of how these steps were performed in this work is presented here.

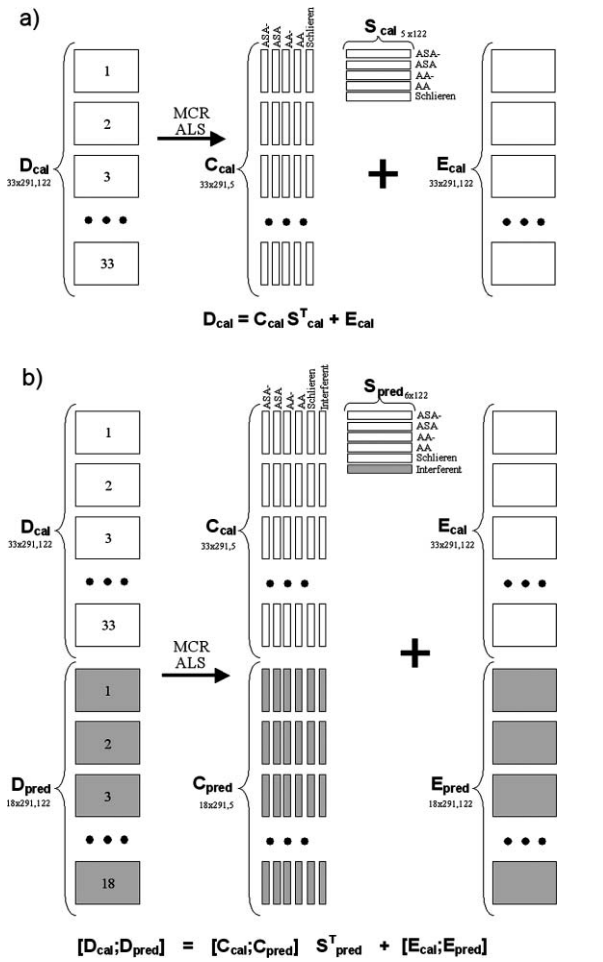

Fig. 1 MCR-ALS analysis of: (a) the calibration column-wise augmented data matrix $\mathbf{D}_{cal}$ composed of 33 synthetic standard calibration samples and (b) the column-wise augmented data matrix $[\mathbf{D}_{cal};\mathbf{D}_{pred}]$ composed of the calibration augmented data matrix ($\mathbf{D}_{cal}$) and the prediction (validation or pharmaceutical samples) column-wise augmented data matrix $\mathbf{D}_{pred}$ composed of 18 prediction samples. $\mathbf{C}_{cal}$ and $\mathbf{C}_{pred}$, are the corresponding column-wise augmented concentration profiles matrices, $\mathbf{S}^{T}_{cal}$ and $\mathbf{S}^{T}_{pred}$ the spectra profiles matrices, and $\mathbf{E}_{cal}$ and $\mathbf{E}_{pred}$ are the column-wise augmented errors matrices and the subscripts "cal" and "pred" mean calibration and prediction (either from validation or from pharmaceutical samples), respectively. Below each figure, the bilinear model equation applied to each case is given.

3.2.1. Determination of the number of components (NC)

Essentially, the number of components for MCR-ALS analysis can be initially estimated in two ways: from the previous knowledge of the investigated system or from the results obtained by singular value decomposition (SVD) of the data matrix of the investigated system. In the first approach, the number of components is estimated from the number of chemical species expected to be present in the sample. For instance in this work, for the simultaneous determination of the mixtures of ASA and AA, a four component model could be proposed, since each analyte produces two species in their acid–base equilibrium. However, at least two additional components would be necessary, since one additional component will be needed

to model the Schlieren effect (in both standards and samples) and another additional component will be needed to model the presence of possible interferences when a pharmaceutical sample is analyzed. SVD analysis can confirm this preliminary knowledge of the system unless rank deficiencies[10,15] and additional interferents are present. It is assumed that in the absence of rank deficiency problems, significant components (either physical or chemical) should give rise to larger singular values than those related to data noise.

3.2.2. Initial estimates

Initial estimates for either the concentration or spectral modes for the previously selected number of components for the MCR-ALS analysis can be obtained by evolving factor analysis (EFA)[16] from methods based on the detection of purest variables in the data set[17] or from evolutionary algorithms (EA).[18–20] In the present work initial estimates were obtained from the spectral profiles estimated by the purest variables detection approach[17]

3.2.3. ALS optimization

The ALS algorithm starts the optimization using either eqn (2) or (3) depending on whether the initial estimates for the concentration $\mathbf{C} = \mathbf{C}_{inic}$, or for spectral profiles $\mathbf{S}^T = \mathbf{S}^T_{inic}$, were used.[13,14]

$$\mathbf{S}^T = \mathbf{C}^+\mathbf{D}^* \quad (2)$$

$$\mathbf{C} = \mathbf{D}^*(\mathbf{S}^T)^+ \quad (3)$$

where $\mathbf{D}^*$ is the data matrix estimated with the selected number of components and "+" indicates the pseudoinverse operation.[21] The optimization procedure iterates the two steps desctibed by eqn (2) and (3) using the previously estimated values in an alternating least squares way under constraints (see below) trying to minimize the residual matrix $\mathbf{E}$ until convergence. In MCR-ALS, the convergence criterion is reached when the relative change in the lack of fit between two consecutive iterations is lower than a threshold value. The lack of fit (LOF) is defined as[13,14]

$$\mathrm{LOF} = 100\sqrt{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J} e_{i,j}^2}{\sum_{i=1}^{I}\sum_{j=1}^{J} d_{i,j}^2}}, \quad e_{i,j} = d_{i,j} - \hat{d}_{i,j} \quad (4)$$

where $d_{i,j}$ is the ith time and jth wavelength experimental data value in matrix $\mathbf{D}$, $\hat{d}_{i,j}$ is its counterpart data value estimated by ALS, and $e_{i,j}$ is its associated residual. If the ALS does not converge, a maximum number of iterations criterion can be used as a stop criterion. In this work 0.1% and 50 iterations were used as a LOF threshold value and as a maximum number of iterations convergence criteria, respectively. These values are usually employed in many applications of MCR-ALS to spectroscopic data (see previous work[5–11]) and they have been shown to be adequate to provide an appropriate fit to the experimental data and residual matrices with intensities close to the noise level.

Fig. 1a illustrates the extension of the MCR bilinear model for the simultaneous analysis of a set of calibration samples giving data matrices numbered $k = 1, 2, \ldots 33$, arranged in the augmented column-wise data matrix $\mathbf{D}_{cal}(33 \times I,J)$. Results of bilinear MCR-ALS decomposition produces the augmented concentration matrix $\mathbf{C}_{cal}(33 \times I,\mathrm{NC})$, the pure spectra matrix $\mathbf{S}^T_{cal}(\mathrm{NC},J)$ and the residuals matrix $\mathbf{E}_{cal}$. Fig. 1b illustrates the same extension of the MCR bilinear model for the simultaneous analysis of calibration $\mathbf{D}_{cal}$, and validation or prediction pharmaceutical drug samples $\mathbf{D}_{pred}$. Results of bilinear MCR-ALS decomposition now produces the new column-wise augmented concentration matrix $[\mathbf{C}_{cal};\mathbf{C}_{pred}](51 \times I,J)$, the new spectra matrix $\mathbf{S}^T_{pred}(\mathrm{NC},J)$ (which contains some common spectra with $\mathbf{S}^T_{cal}$, see below) and the new residuals matrix $[\mathbf{E}_{cal};\mathbf{E}_{pred}]$. In both cases, a column-wise augmented data matrix ($\mathbf{D}_{cal}$ in Fig. 1a and $[\mathbf{D}_{cal};\mathbf{D}_{pred}]$ in Fig. 1b) is set up and a column-wise augmented concentration matrix ($\mathbf{C}_{cal}$ in Fig. 1a and $[\mathbf{C}_{cal};\mathbf{C}_{pred}]$ in Fig. 1b) are obtained where the concentration profiles of each of the resolved components in each sample data matrix are estimated, as well as the matrix of pure spectra profiles of these resolved components, $\mathbf{S}^T_{cal}$ in Fig. 1a and $\mathbf{S}^T_{pred}$ in Fig. 1b. It can be observed that in the later case, and different to the concentration matrix, only one pure spectra matrix is obtained, in which every row defines a single common spectrum profile for every component in the different individual data matrices simultaneously analyzed in the augmented data matrix. Additionally the corresponding unmodelled parts of experimental data are in matrices $\mathbf{E}_{cal}$ and $[\mathbf{E}_{cal};\mathbf{E}_{pred}]$ respectively, which can be used to evaluate the lack of fit and explained variance parameters for different numbers of components.

3.2.4. Constraints

During the ALS optimization, to have physically meaningful solutions and to minimize rotation ambiguities,[13,14] the iterative calculation of $\mathbf{C}$ and $\mathbf{S}^T$ matrices (by eqn (2) and (3)) is subjected to constraints, which can be imposed based either on previous chemical knowledge of the system or on natural restrictions of experimental systems like non-negativity. In this work non-negativity constraints were applied in all cases for both concentration and spectra profiles. Also to avoid scale indeterminacies during ALS optimization, spectra of the resolved components were normalized to norm equal to one for all of them. This constraint does not change the fit and only fixes the scale and avoids computation instabilities. However quantitative interpretation of results should take into account this fact, especially if one data matrix is analyzed. See ref. 5, 6, 13 and 14 for a more detailed description of constraints in MCR-ALS.

When the data set to be analyzed is an augmented data matrix including more than one experiment (see Fig. 1), additional constraints can be included: (1) The pure spectra of the common components (*i.e.* chemical species), present in different standards or samples are equal. As has already been shown (Fig. 1a and 1b) this constraint is already implicit in the bilinear model of the column-wise augmented data matrices; (2) if an analyte, interferent or component is known to be absent in one or more of the individual data matrices simultaneously analyzed and included in the augmented data matrix, this component is considered not to be present during the ALS optimization and its concentration profile is set to be equal to zero. This is especially useful in these cases where some interferents or new components are only present in some of the simultaneously

analyzed data matrices and it is achieved algorithmically in the MCR-ALS program;[22,23] (3) the concentration profiles (their shapes) of the same component or factor in the different individual data matrices included in the augmented data matrix can be forced to be invariant, to have exactly the same shape in all of them (not their intensity or scale, which a change in the different individual concentration matrices). The application of this constraint is closely related with the extension of the bilinear model to the trilinear model for the analysis of three-way data[5,6,13,14] Due to the high reproducibility of the FIA systems and to the nature of the acid–base equilibrium systems under study in this work, the trilinearity constraint is expected to be fulfilled rather well by the experimental data of this work. (4) A final constraint has been used in this work which is especially useful for calibration purposes and quantitative estimations (see below). This constraint is called the 'fixing calibration profiles' constraint and implies that a particular concentration profile or a part of it is fixed to previously known values during the ALS optimizations.

3.3. Extraction and interpretation of quantitative information

After MCR-ALS analysis, relative quantitative information can be directly estimated from the resolved concentration profiles obtained when a set of data matrices from calibration and unknown (prediction) samples are simultaneously analyzed by MCR-ALS (Fig. 1b). Once the concentration profile of a particular component has been properly resolved, the calibration procedure for its quantitative determination can be directly performed in a similar way to the classical univariate way, *i.e.* based on the linear regression between the known concentration of a set of calibration standards of the analyte and the peak areas or heights of their corresponding concentration profiles. After this calibration step, the concentration of the same analyte in unknown samples is predicted by inverse regression of either its peak area or height in the calibration equation previously obtained for the calibration samples. In this work, for this particular application, the concentration of both analytes has been better estimated (more robust) using peak heights of the concentration profiles of the acidic species of the analytes.

Since this calibration procedure used to obtain quantitative information after MCR-ALS analysis is essentially univariate (peak heights of the resolved concentration profiles *versus* concentration of calibration standards), well established figures of merit for calibration procedures were used.[24] In the chemometrics literature a commonly used parameter used for the quantitative agreement between reference and estimated concentration values is based on root mean square error of prediction (RMSEP), defined as:

$$\mathrm{RMSEP} = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{y}_i - y_{\mathrm{ref},i}\right)^2}{N}} \qquad (5)$$

where $y_{\mathrm{ref},i}$ is the reference concentration value for each of the N samples. Apart from this parameter, it is also good to have a relative measure of the error of prediction (REP) independent of the units, which can be obtained as:[25]

$$\mathrm{REP} = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{y}_i - y_{\mathrm{ref},i}\right)^2}{N y_{\mathrm{ref},i}^2}} \qquad (6)$$

To achieve the best resolution and quantification of the analytes in the analysis of the validation samples and of the commercial drugs, MCR-ALS was executed repeatedly using three different strategies (quantification results of analytes using these three strategies will be compared in Table 1 and commented on in the Results section):

1. In a first strategy, only non-negativity and spectra normalization constraints were applied directly as it is shown in Fig. 1b, using the augmented data matrix including both calibration and prediction (validation or pharmaceutical samples) data sets.

2. In a second strategy, non-negativity and spectra normalization constraints were applied together with the trilinearity constraint (see ref. 12 and 13).

3. In a third strategy, after achieving the optimal resolution of the calibration data set as in Fig. 1a, in the sense that the spectra and concentration profiles of the analytes were correctly resolved, the matrix of calibration samples ($\mathbf{D}_{\mathrm{cal}}$) was further augmented with the matrix of prediction samples ($\mathbf{D}_{\mathrm{pred}}$) to give the new augmented data matrix [$\mathbf{D}_{\mathrm{cal}}$;$\mathbf{D}_{\mathrm{pred}}$] and MCR-ALS was executed again, as shown in Fig. 1b. When prediction samples (including validation samples or pharmaceutical tablets) were included in the analysis, the presence of unknown interferents not present in the calibration samples disturbed the optimal resolution of the different components previously achieved for the calibration samples, producing poorer qualitative and quantitative results. In order to prevent this inaccuracy error, the best concentration and spectral profiles ($\mathbf{C}_{\mathrm{cal}}$ and $\mathbf{S}_{\mathrm{cal}}$) previously obtained for the analytes in the analysis of the calibration samples were fixed during the MCR-ALS resolution of the new augmented data matrix which now includes validation or pharmaceutical prediction samples. According to this strategy, only concentration profiles of the components in the prediction samples ($\mathbf{C}_{\mathrm{pred}}$) and spectra of new interferences were allowed to change during the MCR-ALS optimization. This strategy is called 'fixing calibration profiles' constraint in this work. The implementation of this constraint was achieved by using appropriate 'masking' matrices $\mathbf{C}$ and $\mathbf{S}^{\mathrm{T}}$, that update at each ALS iteration the selected matrix elements of $\mathbf{C}$ and $\mathbf{S}^{\mathrm{T}}$ matrices to be equal to the fixed values obtained during the calibration process.

3.4. Software

All the calculations were performed in a MATLAB environment (version 7.0) and the MCR-ALS method was already described in previous references,[10,11,13,14] implemented and downloaded as a toolbox in a set of MATLAB functions available on the Internet[22,23]

4. Results and discussion

4.1. MCR-ALS analysis of FIA-DAD synthetic calibration mixtures of ASA and AA

In Fig. 2 a 3D-plot of the FIA analysis of a standard calibration sample containing 80 and 47 mg L^{-1} of ASA and AA,

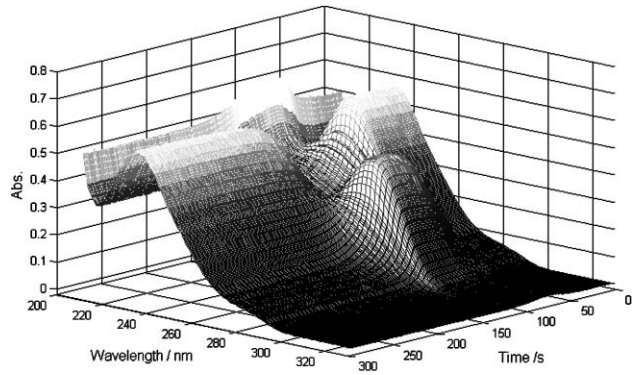

Fig. 2 3D plot of the pH gradient FIA-DAD response of a standard calibration sample containing 80 and 47 mg L^{-1} of ASA and AA respectively.

Table 1 MCR-ALS figures of merit for the results obtained in the analysis of calibration, validation and pharmaceutical samples using different optimization strategies

		RMSEP[a]		REP[a]	
Situation	LOF exp[a]	ASA	AA	ASA	AA
Calibration[b]	0.8291	0.82	0.31	1.18	0.64
Calibration[b,c]	1.9481	0.57	0.37	0.73	0.67
Validation[b]	0.7840	1.95	0.79	3.67	2.03
Validation[b,c]	2.2926	1.71	0.36	2.16	1.04
Validation[b,c,d]	2.1486	1.00	0.38	1.10	1.11
Melhoral[b]	0.7323	4.15	1.63	4.98	2.77
Melhoral[b,c]	2.7509	6.69	2.10	8.08	3.63
Melhoral[b,c,d]	2.5287	6.96	1.91	8.39	3.26
Aspirina[b]	0.7583	1.10	0.58	1.32	1.09
Aspirina[b,c]	2.8773	1.61	1.01	1.99	1.75
Aspirina[b,c,d]	2.8312	1.16	0.92	1.41	1.62
Doril[b]	0.8573	5.34	—	5.80	—
Doril[b,c]	1.8468	0.41	—	0.41	—
Doril[b,c,d]	1.7547	1.47	—	1.60	—
Sandoz[b]	1.2352	—	1.11	—	2.16
Sandoz[b,c]	2.7816	—	0.91	—	1.76
Sandoz[b,c,d]	3.0376	—	1.48	—	2.46

[a] Figures of merit: LOF—lack of fit of experimental data (see eqn (4)); RMSEP/mg L^{-1}—root mean square error in prediction for the MCR-ALS resolved acidic species concentration profile (see eqn (5)); REP (%)—relative error in prediction in percentage for the MCR-ALS resolved acidic species concentration profile (see eqn (6)). [b] Constraint used for MCR-ALS resolution are non-negativity, trilinearity ([c]) and fixing calibration profile ([d]).

respectively, is given. The region where the pH gradient occurs can be clearly distinguished between approximately 100 and 150 s. Under good mixing conditions, with a large dispersion of the two analytes (achieved when a mixing chamber is used in the FIA system) and with an injected sample that does not present a high concentration difference in relation to the carrier stream (as in the present work), the Schlieren effect should give a smooth and relatively low signal.[2] Therefore for this particular FIA system the Schlieren effect was not expected to have a large contribution to the total signal. To verify this and resolve the analyte mixtures, MCR-ALS was applied first to calibration samples composed of the augmented data matrix containing 33 individual data matrices corresponding to 3 replicates of each of the 11 different analyte compositions. Five components, two for each analyte (two acid–base equilibrium species) and one additional component to model the Schlieren effect were considered. Initial estimates for the spectral profiles were obtained from the purest FIA times,[17] *i.e.* spectra at those diffusion FIA times where the different components predominate. In Table 1 (first two rows and second column), results of lack of fit for the MCR-ALS analysis of the FIA-DAD calibration data set in two different cases are given. A first case (first row) is when only non-negativity and spectra normalization constraints were applied and a second case (second row) is when apart from them, also the trilinearity constraint was applied. Lack of fit values (see eqn (4)) were respectively as low as 0.89% and 1.95% (R^2 of explained variances were over 99.9%), which confirms on one hand that the system could be accurately described by the model with 5 components and on the other hand that the system deviates very little from trilinearity. The presence of the fifth species corresponding to the Schlieren effect improved the fit from 2.30% (4 components, not shown in Table 1) to 1.95% (5 components). Although this is only a small improvement, the fact that the shape of the resolved profile for the Schlieren effect (Fig. 3a) was also good, confirms the reliability of the obtained results. This profile shows clearly that the presence of the Schlieren effect is in the gradient zone of each analyzed sample—a time always between 50 and 100 s. It is important to remark here the good resolution of this profile, in spite of its rather low contribution and in spite of its strong overlap with the other four concentrations profiles of the two acid–base forms of the two analytes in the FIA gradient region.

Also in Fig. 3a, the MCR-ALS resolved FIA concentration profiles of the acid and basic species of the two analytes, ASA and AA, are given. The shapes of these resolved concentration profiles are in good agreement with their expected behaviour according to the pH gradient of the FIA system, especially in the transition zones where the composition of the acid–base system changes drastically. In contrast, the absolute heights attained for these concentration profiles are dependent on the spectra normalization constraint used during the MCR-ALS analysis (see Method section) and they do not give directly their true relative concentration. Quantitative information will be derived from the FIA run to run changes when standards with known concentration are simultaneously analyzed (see later in section 4.4). In Fig. 3b the estimated spectra profiles for the two acid–base forms of the two analytes and for the Schlieren effect are shown. The spectrum profile of the Schlieren effect presents a region between 310 and 330 nm where the other components of the system did not absorb, helping the correct resolution of its concentration (diffusion) profile. Another important reason why this resolution was achieved in spite of its intrinsic difficulty, was because of the large number of data matrices simultaneously analyzed (matrix augmentation in Fig. 1a) and of the large number of calibration samples and replicates used in the MCR-ALS analysis (see Experimental section). In Fig. 3c relative heights at the maximum of the MCR-ALS resolved concentration profiles are plotted for both the acid and basic species of the analytes and for the Schlieren effect in the calibration samples data set. Little variation is observed for the heights of the Schlieren in all calibration samples and a much

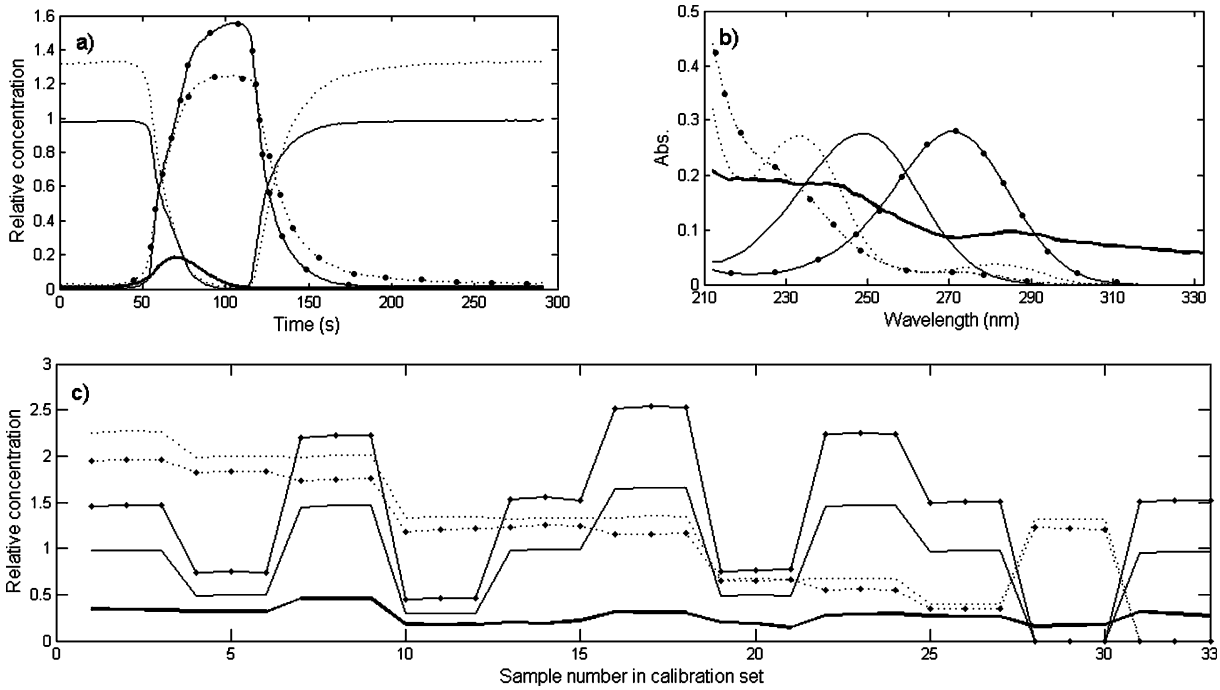

Fig. 3 MCR-ALS estimated component profiles in the analysis of calibration samples ($\mathbf{D}_{cal}$ data matrix): (a) Concentration profiles $\mathbf{C}_{cal}$; (b) spectra profiles $\mathbf{S}^{T}_{cal}$; (c) heights at the maximum of the MCR-ALS resolved concentration profiles for acid and basic species of the analytes and for the Schlieren effect in the calibration samples. ASA (dashed), AA (solid), Schlieren effect (thick solid line), basic species (●).

larger variation is observed for the concentration profiles of the acid and basic species of the analytes in agreement with the variations of the total analyte concentrations stipulated by the experimental design conditions (see Experimental section). It can be observed that the acid–base pair profiles of each analyte follow exactly the same pH pattern for all FIA runs as it should be since both were linked by the acid–base reaction and they both should reach their maximum concentration at the same pH of the FIA gradient system. Quantitative estimations are further discussed later in section 4.4.

4.2. MCR-ALS analysis of FIA-DAD synthetic validation mixtures of ASA and AA

For validation analysis, the new column-wise augmented data matrix was formed using individual data matrices coming from both calibration and validation data sets. This augmented data matrix was composed of 48 matrices, where 33 were from the calibration set and 15 were the new ones from the validation set (see Experimental section). The procedures and constraints used in their analysis by MCR-ALS were the same as the one given below for the commercial pharmaceutical drugs. Due to the additional presence of caffeine as interferent, six components were considered and resolved by MCR-ALS. As shown in Table 1, using the three different strategies previously described in the Method section, lack of fit values for the six components models were in all cases good and within the expected experimental noise level for UV-VIS absorption spectrophotometric measures (around 1–3%). Quantitative results of validation samples are commented on later in section 4.4.

4.3. MCR-ALS analysis of FIA-DAD commercial pharmaceutical drugs

For the analysis of each pharmaceutical compound, a new augmented data matrix was formed using individual data matrices coming both from calibration samples and from pharmaceutical samples as shown in Fig. 1b. This new augmented data matrix was composed of a total number of 51 matrices, where the first 33 matrices were the calibration samples (mixtures of the two analytes at 11 compositions × 3 replicates) and the other 18 matrices were the tablet drug samples (6 tablets × 3 replicates) of each of the four analyzed commercial drugs. Six components were used in this case to model these systems, where now the extra sixth component was added to model some possible interferences present in the tablets. As explained before, in calibration samples this 6th component (interferent) was not present. In Doril® and Sandoz® drugs, AA or ASA were not present respectively. When a particular component is known to be absent in a data matrix, its concentration is kept to zero during the ALS optimization. Initial estimates of the spectral profiles for the protonated (ASA and AA) and unprotonated (ASA^{-} and AA^{-}) species and for the Schlieren effect were taken from the optimal solution obtained in the previous MCR-ALS analysis of the calibration data set.

In Fig. 4a–c MCR-ALS resolved concentration profiles for one of the samples of the pharmaceutical drug Melhoral®, using different constraints (non-negativity, trilinearity and fixing calibration profiles) are shown. It is observed in Fig. 4a that the shapes of the resolved concentration profiles present regular pulses (which are more easily visualized in the ascent of the basic species concentration profile), which can be attributed to the

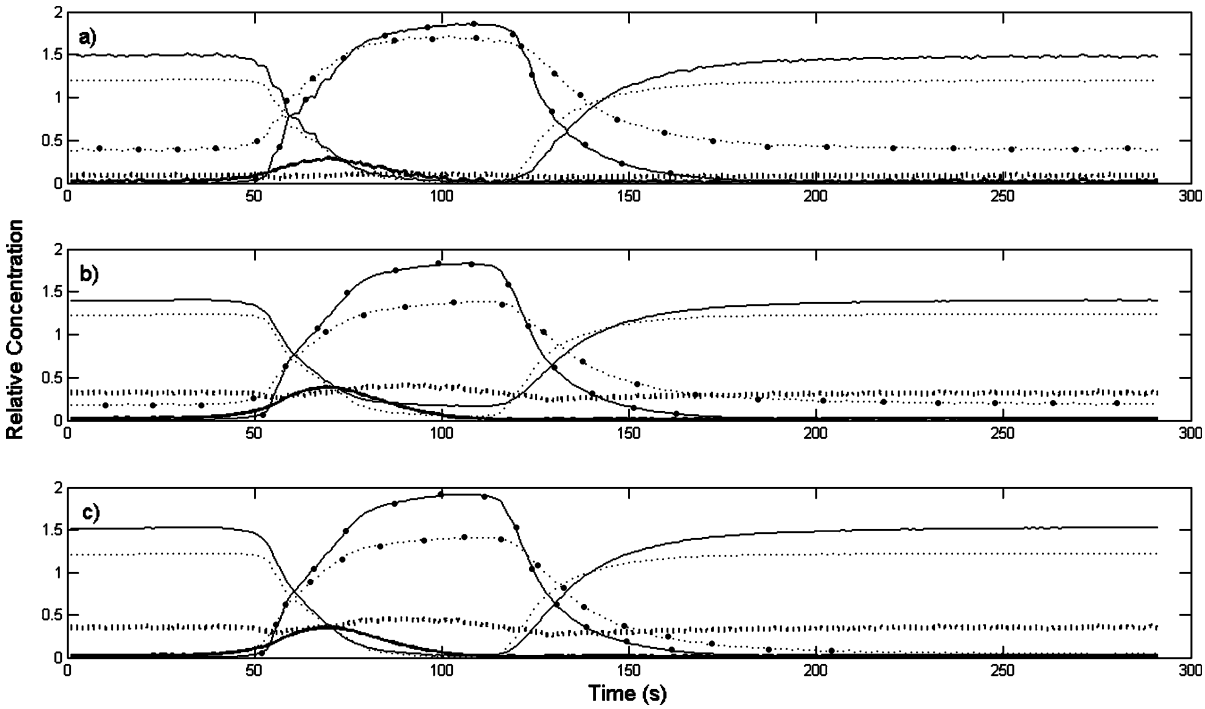

Fig. 4 MCR-ALS resolved concentration profiles for one of the samples of the pharmaceutical drug Melhoral®, using three different optimization strategies: (a) spectra normalization and non-negativity constraints; (b) spectra normalization, non-negativity and trilinearity constraints; (c) spectra normalization, non-negativity, trilinearity and fixing calibration profiles contraints. ASA (dashed), AA (solid), Schlieren effect (thick solid line), interferent (thick dashed line), basic species (●).

peristaltic pump pulses. By contrast, these regular pulses were filtered in the profiles resolved by MCR-ALS when trilinearity and fixing calibration profiles constraints were applied. This explains why the lack of fit increased somewhat (see values in first column of Table 1 for the analysis of this pharmaceutical compound) when these constraints were applied. When we compare the concentration profiles resolved by MCR-ALS under different types of constraint (in Fig. 4a, 4b and 4c, up to down), it is possible to see that the concentration of the basic species of ASA decreases while the concentration of interference increases at the same time. So, the best resolution is observed when the fixing calibration profile constraint (Fig. 4c) is applied, especially for the concentration profile of the basic species of ASA which is expected to be approximately zero before the pH gradient. The same was observed for all tablets that contained ASA. Also the concentration profile resolved for the interferent is practically constant along the FIA experiment, not changing its concentration as a consequence of the pH gradient.

Fig. 5a shows the comparison of spectra profiles obtained experimentally from synthetic samples (containing only one of the acid–base pairs of ASA, AA and caffeine) and of spectra profiles recovered by MCR-ALS in the simultaneous analysis of calibration and pharmaceutical samples for Doril® drug, using the constraints of non-negativity, trilinearity and fixing calibration profiles, as explained above. MCR-ALS resolved spectrum for caffeine was obtained as the interference spectrum profile in the Doril® drug, when the three constraints (non-negativity, trilinearity and fixing calibration profiles) were used. All MCR-ALS resolved spectra were in very good agreement with experimental spectra for both acid and basic species of the two analytes and for caffeine. Finally, Fig. 5b shows the concentration profiles obtained for the Doril® drug. For clearness, just the basic species of ASA (ASA^-), Schlieren effect and the interferent caffeine has been plotted in the figure. In this plot, it can be clearly noticed that it is only between pharmaceutical samples 34 to 51 that the caffeine concentration profile had a constant signal over time, which is in agreement with what was expected since caffeine was only present at constant concentration in these drugs and its concentration profile should not change with the FIA pH gradient.

4.4. Quantification results

Table 1 (columns 3, 4, 5 and 6) summarizes the quantitative results obtained by MCR-ALS for the calibration and validation sets, and for the four pharmaceutical products, with the different constraints discussed above. The results shown were obtained considering the maximum height achieved for the acid species concentration profile of both ASA and AA, which gave the best RMSEP and REP values (when they were compared with the concentrations obtained independently by reference methods, see Experimental section). For calibration synthetic mixtures, no significant differences were observed in RMSEP and REP values when non-negativity and trilinearity constraints were applied (two first rows in Table 1). REP% values did not increase significantly with values of 0.73% and 0.67% (with trilinearity constraint), and of 1.18% and 0.64% (without trilinearity constraint) for five components, for ASA and AA respectively. This proves that the system fulfills rather well the

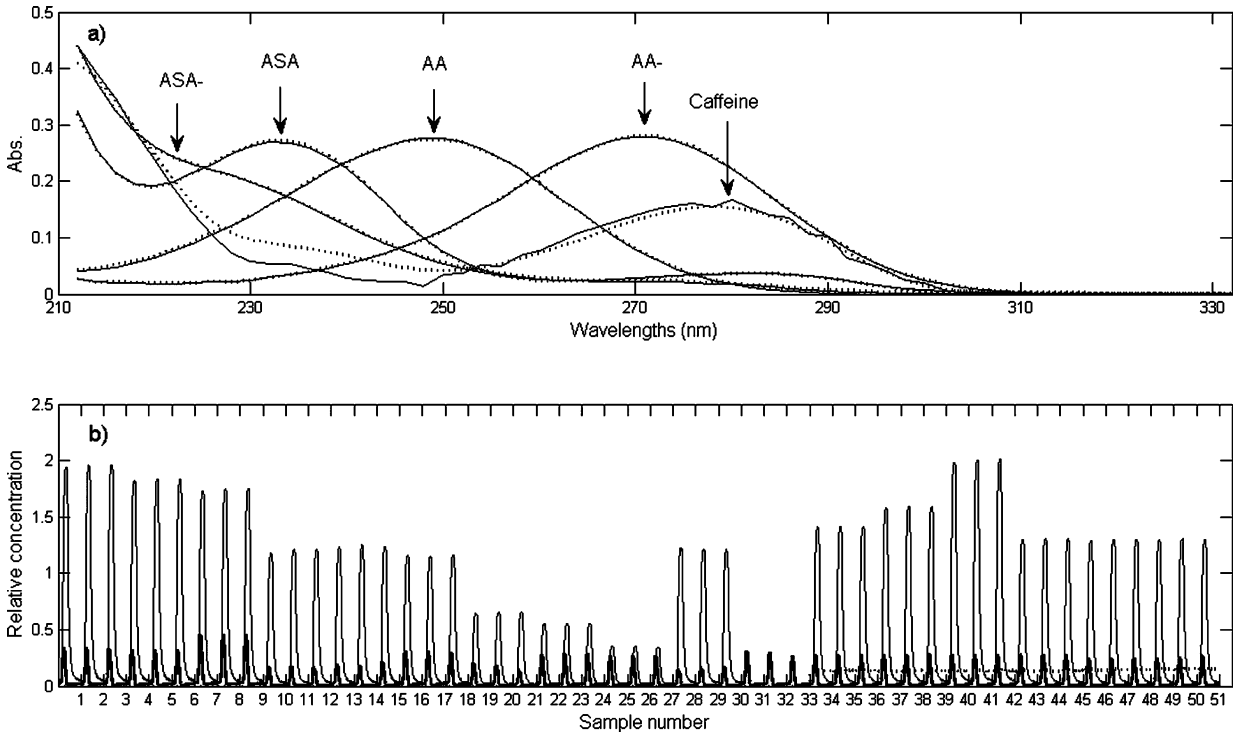

Fig. 5 (a) Comparison of spectra profiles obtained experimentally (dashed) with spectra profiles (solid) recovered by MCR-ALS in the simultaneous analysis of calibration and pharmaceutical samples for the Doril® drug, using the constraints of non-negativity, trilinearity and fixing calibration profiles. (b) MCR-ALS resolved concentration profiles in the simultaneous analysis of calibration and prediction samples [$\mathbf{D}_{cal}$;$\mathbf{D}_{pred}$]: basic ASA^- species (solid), Schlieren effect (thick solid line) and caffeine (dashed).

trilinear model (the shape and FIA-pH gradient position of the concentration profiles of the different components of the system are practically the same for all the simultaneously analyzed calibration samples).

For the validation synthetic mixture samples the best results were obtained when fixing calibration profiles constraint was applied (REP% values of 1.10% and 1.11% for ASA and AA respectively in the 3rd and 4th row of Table 1). MCR-ALS recovered spectra and concentration profiles for caffeine in these samples were also good and in good agreement with those obtained for the interferent in the Doril® drug (already shown in Fig. 5).

In the case of pharmaceutical tablet samples (rows 6 to 17 in Table 1), in almost all cases results showed some increase on prediction errors when trilinearity or both trilinearity and fixing calibration profile constraints were imposed on the model. This increase can be attributed to some changes in the shape and/or position of the concentration profiles, which could be interpreted as a loss of the trilinearity condition. However, based on a F-test with 18 and 18 degrees of freedom, performed with RMSEP values, differences in average prediction errors using the different constraints presented in Table 1 were not significant and therefore, results were not totally conclusive yet. Results for LOF (2nd column in Table 1), showed also a small increase when more constraints were imposed on the model, which can be partially attributed to the elimination of the pulses from peristaltic pumps and to small random variations of time/pH relations in the gradient. However, since these differences were not significant in the predicted concentration values, and since trilinearity and fixing calibration profile constraints gave better shapes of the MCR-ALS resolved concentration profiles estimated independently for each of the investigated chemical system, this approach was the one finally considered to be the best one. Except for ASA in the Melhoral® drug, in all cases REP% values were below 5%. In the case of Melhoral®, REP% values for ASA increased up to 8.4% (when trilinearity and fixing concentration profiles were applied). A possible explanation for that is that Melhoral is a big effervescent tablet and has flavoring and coloring. It is possible then that these poorer results are due to the presence of an additional unmodelled interference with ASA, since for AA good results were obtained. However, no further investigation was carried out to correct this possible inaccuracy.

5. Conclusions

MCR-ALS with the fixing calibration profile constraint was proven to provide good resolutions of the pH gradient FIA systems used in this work for the determination of AA and ASA in pharmaceutical tablets. Quantitative results were in good agreement with independent reference methods. REP errors were approximately 8% for Melhoral® with ASA, and lower than 5% for the other three pharmaceutical products, Aspirina® +C, Doril® and Sandoz®, for both ASA and AA, which confirms that the proposed pH gradient FIA system combined with MCR-ALS can be used as a good alternative method for the simultaneous determination of both analytes in pharmaceutical samples.

In this particular application, due to good mixing conditions, large dispersion properties and compatibility among concentrations of reagents and analytes in the sample, the Schlieren effect only represented a small part of the total signal. However, even representing just a little part of the FIA signal and being highly overlapped with the other four concentration profiles of the acid–basic pairs of ASA and AA, the proposed MCR-ALS method could resolve adequately the contribution of the Schlieren effect as an independent component. This suggests that MCR-ALS can be a good and effective way to model the contribution of this effect in general and that it allows for the correction of this physical phenomenon in the FIA-DAD systems with pH gradient, like the one studied in this work.

Acknowledgements

The authors acknowledge financial support and fellowships from the UNICAMP PED-A Program, from FAPESP (proc. 05/53280–4) and Santander/Banespa Bank for an international mobility scholarship.

References

1 J. Ruzicka and E. H. Hansen, in *Flow Injection Analysis*, John Wiley & Sons, New York, 2nd edn, 1988.
2 A. C. Dias, E. P. Borges, E. A. G. Zagatto and P. J. Worsfold, *Talanta*, 2006, **68**, 1076.
3 F. J. Krug, H. Bergamin, E. A. G. Zagatto and S. S. Joergensen, *Analyst*, 1977, **102**, 503.
4 K. Booksh and B. R. Kowalski, *Anal. Chem.*, 1994, **66**, 782A.
5 J. Diework, A. de Juan, R. Tauler and B. Lendl, *Appl. Spectrosc.*, 2002, **56**, 40.
6 J. Diework, A. de Juan, M. Maeder, R. Tauler and B. Lendl, *Anal. Chem.*, 2003, **75**, 641.
7 J. Saurina, S. Hernández-Cassou and R. Tauler, *Anal. Chim. Acta*, 1996, **335**, 41.
8 J. Saurina, S. Hernández-Cassou and R. Tauler, *Anal. Chem.*, 1997, **69**, 2329.
9 J. Saurina, S. Hernández-Cassou, R. Tauler and A. Izquierdo-Ridorsa, *Anal. Chem.*, 1999, **71**, 2215.
10 A. Izquierdo-Ridorsa, J. Saw-ha, S. Hernández-Cassou and R. Tauler, *Chem. Intell. Lab. Syst.*, 1997, **38**, 183.
11 J. Saurina, S. Hernández-Cassou, R. Tauler and A. Izquierdo-Ridorsa, *Anal. Chim. Acta*, 2000, **408**, 135.
12 *Official Monographs*, United States Pharmacopoeia USP28NF23, 2005, p. 178.
13 R. Tauler, A. Smilde and B. Kowalski, *J. Chemom.*, 1995, **9**, 31.
14 R. Tauler, *Chem. Intell. Lab. Syst.*, 1995, **38**, 133.
15 J. Saurina, S. Hernández-Cassou, R. Tauler and A. Izquierdo-Ridorsa, *J. Chemom.*, 1998, **12**, 183.
16 M. Maeder and A. D. Zuberbüehler, *Anal. Chim. Acta*, 1986, **181**, 287.
17 W. Windig and J. Guilment, *Anal. Chem.*, 1991, **63**, 1425.
18 C. R. Houck, J. A. Joines and M. G. Kay, *Evol. Comput.*, 1997, **5**, 31.
19 C. R. Houck, J. A. Joines and M. G. Kay, *Comput. Oper. Res.*, 1996, **23**, 587.
20 A. Kollen and E. Pesch, *Discrete Appl. Math. Comb. Oper. Res. Comput. Sci.*, 1994, **48**, 273.
21 G. H. Golub and C. F. Van Loan, in *Matrix Computations*, John Hopkins Series in Mathematical Sciences, London, 2nd edn, 1989.
22 http://www.ub.es/gesq/mcr/mcr.htm, accessed in January 2007.
23 J. Jaumot, R. Gargallo, A. de Juan and R. Tauler, *Chem. Intell. Lab. Syst.*, 2005, **76**(1), 101.
24 J. Saurina, C. Leal, R. Compañó, M. Granados, M. D. Prat and R. Tauler, *Anal. Chim. Acta*, 2001, **432**, 241.
25 N. J. Miller-Ihli, T. C. O'Haver and J. M. Harnly, *Spectrochim. Acta, Part B*, 1984, **39**, 1603.