Background Correction and Multivariate Curve Resolution of Online Liquid Chromatography with Infrared Spectrometric Detection

ARTICLE *in* ANALYTICAL CHEMISTRY · JUNE 2011

Impact Factor: 5.64 · DOI: 10.1021/ac2004407 · Source: PubMed

5 AUTHORS, INCLUDING:

Julia Kuligowski
Instituto de Investigación Sanitaria La Fe
60 PUBLICATIONS **348** CITATIONS
SEE PROFILE

Guillermo Quintas
Leitat Technological Center
80 PUBLICATIONS **835** CITATIONS
SEE PROFILE

Bernhard Lendl
TU Wien
270 PUBLICATIONS **4,315** CITATIONS
SEE PROFILE

Miguel de la Guardia
University of Valencia
570 PUBLICATIONS **9,534** CITATIONS
SEE PROFILE

analytical chemistry

ARTICLE

pubs.acs.org/ac

Background Correction and Multivariate Curve Resolution of Online Liquid Chromatography with Infrared Spectrometric Detection

Julia Kuligowski,[†] Guillermo Quintás,[*,‡] Romà Tauler,[§] Bernhard Lendl,[⊥] and Miguel de la Guardia[†]

[†]Department of Analytical Chemistry, University of Valencia, Edificio Jerónimo Muñoz, 50th Dr. Moliner, E-46100 Burjassot, Spain

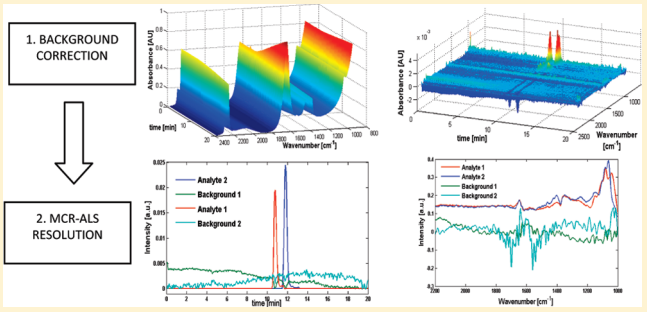
[‡]Bio InVitro Division, Leitat Technological Center, de la Innovacio 2, E-08225 Terrassa, Spain

[§]Environmental Chemometrics Group, Department of Environmental Chemistry, Institute for Chemical and Environmental Research, CID-CSIC, Jordi Girona 18, E-08034 Barcelona, Spain

[⊥]Institute of Chemical Technologies and Analytics, Vienna University of Technology, Getreidemarkt 9/164, A-1060 Vienna, Austria

Supporting Information

ABSTRACT: The use of multivariate curve resolution—alternating least-squares (MCR-ALS) in liquid chromatography—infrared detection (LC-IR) is troublesome due to the intense background absorption changes during gradient elution. Its use has been facilitated by previous removal of a significant part of the solvent background IR contributions due to common mobile phase systems employed during reversed phase gradient applications. Two straightforward background correction approaches based on simple-to-use interactive self-modeling mixture analysis (SIMPLISMA) and principal component analysis (PCA) are proposed and evaluated on reversed phase gradient LC-IR data sets obtained during the analysis of carbohydrate and nitrophenol mixtures. After subtraction of the calculated background signal, MCR-ALS provided improved signal-to-noise ratios, removed remaining mobile phase and background signal contributions, and resolved overlapping chromatographic peaks. The present approach tends to enable easy-to-use background correction to facilitate the use of MCR-ALS in online LC-IR, even in challenging situations when gradient conditions are employed and only poor chromatographic resolution is achieved. It, therefore, shows great potential to facilitate the full exploitation of the advantages of simultaneous quantification and identification of a vast amount of analytes employing online IR detection, making new exciting applications more accessible.



Infrared (IR) spectroscopy is known to be a versatile technique for the detection of a vast amount of molecules, enabling qualitative and quantitative analysis. For this reason, IR detection in liquid chromatography (LC) is an interesting alternative to other detection systems, providing a more selective response than, for example, evaporative light scattering (ELS), refractive index (RI), UV, or electrochemical detectors. In spite of instrumental developments in solvent removal interfaces for off-line[1] detection, online liquid chromatography—infrared detection (LC-IR) hyphenation is still the most straightforward approach because of its easy implementation. This topic has recently been reviewed.[2,3]

Recent technological advances in the field of IR light sources and detectors[4,5] as well as miniaturized systems[6] (e.g., micromachined nL flow cell) promise to significantly enhance the sensitivity of online LC-IR, which especially employing reversed phase conditions is still a barrier for many real life analytical applications.[7] Despite late technical advances, the applicability of online hyphenation is hindered by changes in the mobile phase composition and the presence of intermolecular interactions between the mobile phase constituents that lead to strong changes in shape and intensity of the IR background absorption during gradient LC runs. Although a number of methods for background correction in online LC-IR have been proposed, automatic background correction is still an issue of utmost importance in LC-IR chromatographic systems aimed to work in a fully unattended way.

Using a cubic smoothing splines (CSS) based approach[8] for background correction in LC-IR, spectra measured in the region before and after chromatographic peak clusters are used as knots to model the variation of the eluent absorption intensity with time. The correct selection of the knots is critical to recover true peak heights and shapes, and it can be troublesome when the peaks are not completely baseline resolved, as it frequently occurs in the analysis of complex samples. An alternative strategy based on the use of a reference-spectra matrix (RSM)[9] has also been proposed.

Received: February 27, 2011
Accepted: May 4, 2011
Published: May 04, 2011

It uses the spectral information content obtained during the equilibration of the LC system to perform background correction during the LC-IR run. On the basis of this scheme, a number of approaches have already been developed and evaluated.[3] These methods present, however, three main limitations: (i) high instrument stability is required, (ii) the mobile phase spectra should contain a spectral region free from interferences of the eluting compounds which must be characteristic for its composition, and (iii) the number of spectra included in the RSM affects the accuracy of the correction.

Hyphenated LC instruments are capable of providing two-dimensional bilinear data, and thus, factor analysis (FA) based methods have been employed to explore the underlying sources of variation in this type of data. Among these techniques, principal component analysis (PCA),[10] evolving factor analysis (EFA),[11,12] and simple-to-use interactive self-modeling analysis (SIMPLISMA)[13] have been widely used for the selection of pure spectra and elution profiles, for the determination of the number of eluting components with linearly independent spectra and/or eluting profiles (i.e., the *chemical rank*, which can be estimated by the pseudorank or by the mathematical rank in absence of noise[14]) and for the resolution of overlapping or coeluting peaks. Besides, methods such as parallel factor analysis (PARAFAC,[15] PARAFAC2[16,17]) or multivariate curve resolution−alternating least-squares (MCR-ALS)[18−23] have been used to resolve the pure concentration profiles and pure instrumental responses in a variety of analytical situations. An important feature of all these methods is their second order advantage, which allows for the quantification of the target analytes in the presence of unknown interferences when several chromatographic runs are analyzed together.

In 2003, Edelmann et al.[24] employed MCR-ALS for the quantitative analysis of coeluting analytes in wine samples using isocratic LC-Fourier transform infrared spectrometry (LC-FT-IR). In this work, prior to MCR-ALS, background correction was achieved by subtracting the average baseline spectrum recorded directly before coelution of the analytes and calculating first-derivative spectra in order to remove baseline drifts. Ruckebusch et al.[25] used MCR-ALS for resolving gel permeation chromatography-FT-IR (GPC-FT-IR) data collected on butadiene rubber and styrene butadiene rubber blends in order to access in-depth knowledge of polymers along the molecular weight distribution.

In 2004, Boelens et al.[26,27] developed a very efficient method to correct for the eluent background spectrum (EBS) in LC coupled to spectroscopic techniques (e.g., Raman, UV−vis). In the EBS method, data acquired during the LC run is split into two parts. One part is a matrix where only the eluent is present (i.e., the background spectral subspace or B-subspace), and the other part is formed by spectra in which eluent and analyte(s) are both present (i.e., the elution spectra). First, all variation in the eluent spectra at baseline level is modeled in the background spectral subspace (i.e., B-subspace) built by PCA using a number of principal components selected according to the empirical IND algorithm.[28] Second, the PCA loading vectors are used for eluent background correction of the elution spectra. Therefore, in this approach, it is assumed that the elution spectra are a linear combination of analyte and background spectra, latter being described by the previously calculated loading vectors. These are fitted under the elution spectra by an asymmetric least-squares (asLS) method.[29] This algorithm assumes that the analyte spectrum consists of positive values only and negative residuals of the least-squares fit are penalized while performing the regression in an iterative way. In our opinion, the main limitation of the EBS method, as described by Dijkstra et al.,[27] is that the elution time window of the analytes has to be known in advance, thus limiting the applicability of the approach. For instance, if the identification of the elution window of the analytes is performed using an additional detection system like, for example, UV, the presence of non-UV absorbing analytes might lead to an underestimation of the number of PCs required to describe the background absorption of the "analyte" subspace. Besides, as the data set has to be split prior to background correction, it cannot be carried out on-the-fly.

In 2006, the use of PARAFAC and PARAFAC2 for chemometric eluent elimination was evaluated by István et al.[30] From results found on simulated online LC-IR runs, the authors concluded that PARAFAC2 performs better than PARAFAC, but like MCR-ALS, it did not give correct decompositions. The use of a new method named objective subtraction of solvent spectrum with iterative use of PARAFAC and PARAFAC2 (OSSS−IU−PARAFAC and OSSS−IU−PARAFAC2, respectively) improved results. In spite of that, the restrictions imposed by this approach (e.g., constant eluent composition or constant elution profiles of any given component among different LC runs) drastically reduced its practical applicability.

In summary, there is an ongoing interest in compensating changes in baseline contributions in LC-IR. Besides, the resolution process of LC-IR data could be improved if solvent contribution, drift, and detector changes are removed in advance.[20] Although different approaches have already been developed to overcome the challenges of online (gradient) LC-IR, there is still a need to develop user-friendly strategies with fewer conditions of applicability. Accordingly, the goal of the present study was to develop new approaches for background correction using the information content of a RSM and, hence, facilitate the use of MCR-ALS in the analysis of hyphenated LC-IR data. In summary, two problems are studied in this work: (i) the development and application of background elimination methods, and (ii) the use of MCR-ALS in online gradient LC-IR to resolve coelution problems after background elimination and to improve background correction. This paper also discusses several practical issues encountered in background correction in online LC-IR systems with the help of real gradient LC-IR data sets.

THEORETICAL BACKGROUND

Background Correction. Data obtained from hyphenated LC systems can be seen as the sum of background, noise, and analyte contributions.[31] Background and analyte contributions can be considered as bilinear contributions of spectra and concentration profiles. This implies that background absorption can be modeled as a combination of the contributions of a series of "pure" mobile phase components with their characteristic spectra and with their contribution to the overall absorption calculated according to their concentration or elution profile during the chromatographic run.

Accordingly, the methods described in this work are based on the calculation of both, the spectra and concentration profiles of the IR-absorbing components of the mobile phase during the chromatogram. These methods can be summarized in four steps: (i) Splitting of the LC-IR data matrix in two new matrices, **SM** (n x j) and **RSM** (m x j), corresponding to the spectra acquired during the sample elution (n) and during the re-equilibration

after the LC run (m), respectively, at j wavenumbers. Therefore, whereas the **RSM** matrix contains information only related to the spectral changes due to the mobile phase gradient, the **SM** also includes spectral variations due to the elution of the IR absorbing analytes. (ii) Determination of the number of chemical components due to the mobile phase gradient (k_{opt}) in the **RSM** employing factor analysis. (iii) Calculation of the concentration profiles of the k_{opt} background components in the **SM** matrix, and (iv) calculation and subtraction of the background contribution in the **SM** matrix. Following this scheme, two methods are described in this work.

Method 1: Background Correction Using SIMPLISMA. In this method, the spectra of a user-defined number of "pure" components (k_{opt}) are calculated from the **RSM** data matrix using SIMPLISMA. A critical step is the determination of the chemical rank (i.e., k_{opt}), as the over or under estimation of that value will impair eluent subtraction accuracy. The use of the log(*eigenvalues*) obtained from a singular value decomposition (SVD) analysis[19] of the **RSM** matrix is proposed. The chemical rank has a good correlation with the number of significant singular values of the data matrix, providing that the spectra of the absorbing components are linearly independent of each other.[32] After determination of the chemical rank, the spectra of the k_{opt} "pure" components (i.e., $\mathbf{S_k}$ ($j \times k_{opt}$)) are calculated from the **RSM** data matrix using SIMPLISMA.

Each spectrum r ($1 \times j$) included in the **SM** ($n \times j$) can be expressed by the following equation:

$$r = \mathbf{C_k}\mathbf{S_k}^{\mathrm{T}} + \mathbf{Z} \quad (1)$$

where $\mathbf{C_k}$ ($1 \times k_{opt}$) corresponds to the concentrations of the k_{opt} background components with pure spectra $\mathbf{S_k}$ and $\mathbf{Z}$ ($1 \times j$) corresponds to the absorbance due to the eluting analytes.

As the background absorption is much more intense than that due to the analytes, for each spectrum r to be background corrected, the corresponding $\mathbf{C_k}$ values can be directly estimated by classical least-squares (CLS) using eq 2:

$$\mathbf{C_k} = (\mathbf{S_k}^{\mathrm{T}}\mathbf{S_k})^{-1}\mathbf{S_k}^{\mathrm{T}}r^{\mathrm{T}} \quad (2)$$

Afterward, the background-corrected chromatogram ($\mathbf{Z}$) is obtained by subtracting the estimated spectral background.

Using the logarithm of the *eigenvalues* for chemical rank selection is simple, but it might fail in some cases, like in the presence of significant components at low concentrations.[33] Therefore, in addition, the optimization of the background correction accuracy of the **RSM** matrix as a function of k_{opt} is proposed. Using venetian blinds cross validation with one data split, the **RSM** ($m \times j$) is split into two blocks: **A** and **B**. Then, the SIMPLISMA background correction approach described before is performed iteratively on **B** employing the spectra of k components ($k = (1,\ldots, k_{max})$) extracted using the **A** subdata set as **RSM**. For each considered k value, the background correction accuracy is measured as the mean absolute values of the corrected matrix **B**. The same process is performed using **B** as "reference matrix" for the background correction of **A**. Finally, the mean of the correction accuracy values for each k value is calculated. From results obtained, the k_{opt} is selected as a function of the mean correction error which has to drop down to a threshold value defined by the user based on previous experience (e.g., noise values obtained under isocratic conditions using the same LC-IR set up).

Method 2: Background Correction Using PCA. First, background absorption during gradient elution is gathered in the k_{opt} loading vectors **P** ($j \times k_{opt}$) calculated by PCA of the spectral subspace formed by the **RSM** data set. Then, new observations (i.e., spectra in **SM**) are projected onto the hyperspace defined by the PCA loading vectors to obtain their scores (**T** ($n \times k_{opt}$)). Next, the background estimate of the **SM** matrix is calculated as $\mathbf{TP}^{\mathrm{T}}$. Afterward, the background-corrected chromatogram is calculated by subtracting the estimated spectral background from the raw data.

The calculation of the Q-statistic as the sum of squares of the residual values of the new samples (i.e., **SM**) provides a measure of the variation of the new data outside of the PCs included in the model[34,35] calculated from the **RSM** data. Therefore, if the presence of analytes in the **SM** spectra produces a change in the covariance structure of the model, the increase in the corresponding Q values obtained would allow for an easy identification of the elution windows of the analytes. This information can also be used, for example, to include "background" **SM** spectra for a postrun background correction. The use of the **RSM** data is not limited to the background correction methods described in this work. For example, the use of the **RSM** matrix as B-subspace in the EBS method[26,27] appears to be a promising way to facilitate the use of this method as well.

Multivariate Curve Resolution−Alternating Least-Squares (MCR-ALS). MCR-ALS is a widely used multivariate self-modeling curve resolution.[19] When multivariate hyphenated chromatographic data is analyzed, every chromatographic run can be expressed as:

$$\mathbf{D} = \mathbf{CS}^{\mathrm{T}} + \mathbf{E} \quad (3)$$

where rows of matrix **D** ($n \times j$) are the n spectra recorded and columns are the elution profiles at the j wavenumbers, **C** ($n \times k$) is the matrix of the elution profiles of the k compounds resolved, $\mathbf{S}^{\mathrm{T}}$ ($k \times j$) is the matrix of their corresponding resolved spectra, and **E** ($n \times j$) corresponds to unexplained data variance.

To solve the bilinear model given before, an ALS algorithm is used.[20] The iterative ALS algorithm calculates at each iteration estimates of **C** and $\mathbf{S}^{\mathrm{T}}$ pure matrices optimizing the fit to the original data matrix **D** and minimizing the remaining unexplained variance **E**. The optimization is performed for a number of components and employs initial estimates of either **C** or $\mathbf{S}^{\mathrm{T}}$ obtained using either evolving factor analysis (EFA) or SIMPLISMA, among other possible methods. The optimization is finished when the restitution of the original data matrix **D** is satisfactory and a convergence criterion is met.[18,33] ALS results can be improved when constraints are applied during the optimization with the goal of obtaining calculated spectra and concentration profiles with physically sound shapes[18,19] like nonnegative spectra and concentration profiles and unimodal elution profiles.

MATERIAL AND METHODS

Reagents and Instruments. Data set 1 consists of FT-IR spectra collected during a reversed phase gradient LC injection of a diluted soft drink sample[36] using a linear CH_3CN/H_2O gradient from 75 to 55% (v/v) CH_3CN. Data set 2 consists FT-IR spectra collected during the reversed phase gradient LC injection of a four nitrophenols standard mixture[37] running mobile phase gradients from 65:35 (v/v) H_2O (0.08% v/v TFA)/CH_3CN (0.08% v/v TFA) to 15:85 (v/v) H_2O (0.08%

dx.doi.org/10.1021/ac2004407 |*Anal. Chem.* 2011, 83, 4855–4862

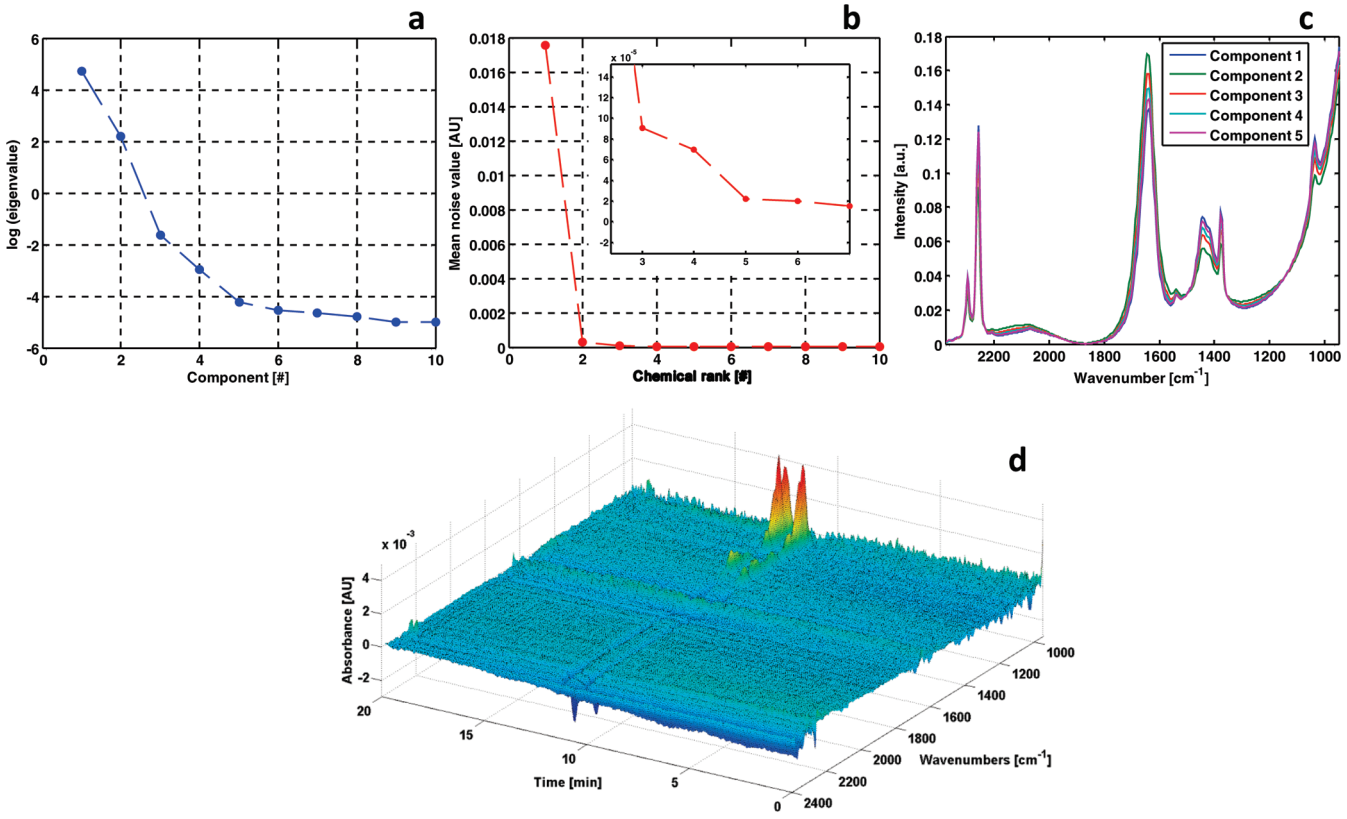


Figure 1. Results of the background correction of the data set 1 following the SIMPLISMA approach. See text for details.

v/v TFA)/CH_3CN (0.08% v/v TFA). RSM for both examples were recorded during system equilibration after gradient injection. Further experimental details are included in the Supporting Information section.

Software and Algorithms. Bold-face capital letters represent matrices; bold-face lowercase characters represent vectors, and italic, lower case letters represent scalars. Online LC-IR data was collected in spectral data matrices $\mathbf{X}$ ($n \times j$) where each row contains a spectrum measured at j wavenumbers. For data acquisition as well as instrumental and measurement control, the OPUS software (version 6.5) from Bruker was employed. Data analysis was run under Matlab 7.7.0 from Mathworks (Natick, USA, 2004) using in-house written MATLAB functions available from the authors and the PLS Toolbox 6.0 from Eigenvector Research Inc. (Wenatchee, WA, USA). Multivariate Curve Resolution was carried out using the MATLAB MCR-ALS toolbox[38] available at http://www.mcrals.info/.

Prior to background correction, negative absorbance values were removed in each spectrum by subtraction of its minimum value. The quality of the background corrected and MCR-ALS recovered spectral profiles was evaluated using the criterion of similarity ($0 \leq R \leq 1$),[39] which entails a comparison of reference spectra with the spectra to be evaluated according to the following equation:

$$R = \frac{\mathbf{r}_i^T \mathbf{s}_i}{||\mathbf{r}_i|| \; ||\mathbf{s}_i||} \qquad (4)$$

where $\mathbf{r}_i$ and $\mathbf{s}_i$ are the reference and recovered spectra of the compound i.

RESULTS AND DISCUSSION

Data Set 1: Carbohydrate Analysis. LC-IR spectra included in data set 1 acquired during an acetonitrile/water LC gradient injection are depicted in Figure S-1 (Supporting Information). In reversed phase LC, the eluent absorbs intensely in the mid-IR region, obscuring the absorption of the analytes.[40] In this type of measurement, the intensities of the eluting analytes are usually <10 mAU, whereas the mobile phase signal reaches intensities up to 1000 mAU. Strong changes in intensity and shape of the background signal can clearly be observed in the region from 2400 to 950 cm^{-1}, where most of the investigated analytes show their characteristic bands. A detailed view of the H_2O bending vibration in the 1640 cm^{-1} region is provided to stress the occurring band shifts during the gradient which complicate accurate background compensation. Detailed description of the IR bands in CH_3CN/H_2O systems can be found elsewhere.[41]

Background Correction. Figure 1 displays results obtained using the SIMPLISMA background correction approach. Figure 1a shows the plot of the log(*eigenvalues*) in descending order obtained from the SVD analysis of the **RSM** (456 × 372) matrix. This figure suggests the use of five components for background correction. The selection of this k_{opt} was supported by results depicted in Figure 1b. In this figure, the effect of increasing k_{opt} on the background accuracy is evaluated using the **RSM** and Venetian blinds cross validation (CV) with one data split. The spectral noise values depicted in Figure 1b do have a meaning in the spectroscopic sense, since they are in the same units as the absorbance measurements and they can be compared directly with known estimates of noise level facilitating the selection of the number of components. In this case, for

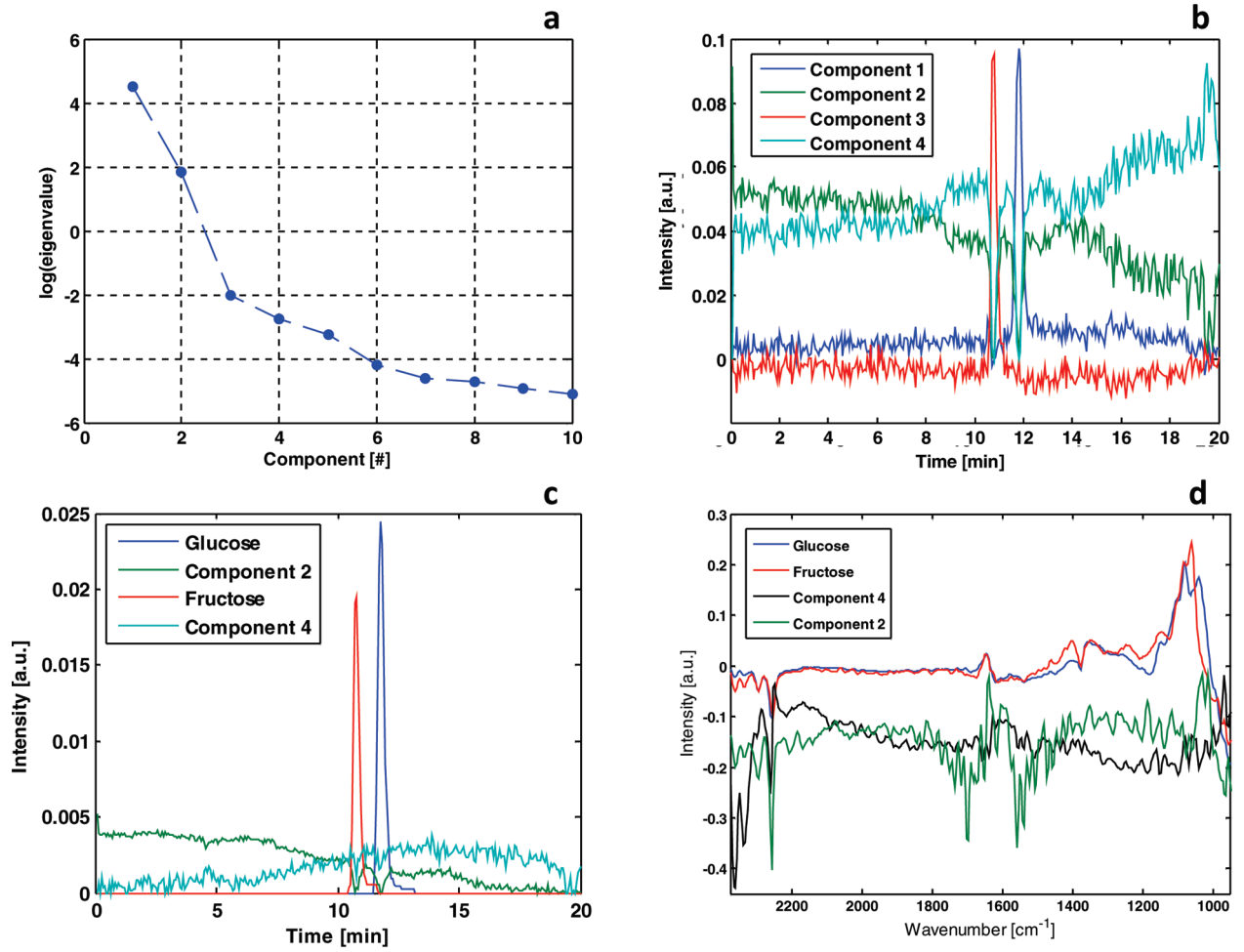


Figure 2. MCR-ALS analysis of the data set 1 after background correction: (a) selection of the chemical rank using SVD; (b) SIMPLISMA initial estimates of the concentration profiles; (c) calculated MCR-ALS concentration traces, and (d) calculated MCR-ALS spectra.

5 components, the mean noise level was 2×10^{-5}, which agrees with our independent estimations of noise level (1.85×10^{-5}) under isocratic conditions. The combined use of both plots is an effective way to simplify the k_{opt} selection process.

Figure 1c shows the corresponding spectra of the pure components obtained using SIMPLISMA (i.e., $\mathbf{S_k}$). Band shapes and relative intensities of the extracted spectra are similar and describe the changes in intensity and shape observed during the CH_3CN/H_2O gradient shown in Figure S-1 (Supporting Information). Using these solvent spectra estimates, their corresponding elution profiles during the sample gradient (i.e., $\mathbf{C_k}$) were calculated by ordinary classical least-squares, CLS. Subsequently, the whole solvent background contribution obtained by calculating the product of $\mathbf{C_k S_k}^T$ was subtracted from the **SM** matrix, obtaining data shown in Figure 1d. From the analysis, fructose and glucose could be identified in the analyzed soft drink sample. The correlation coefficients between reference analyte spectra and those extracted at the peak apex after background correction from 1396 to 1022 cm^{-1} were 0.95 and 0.88, for fructose and glucose, respectively. As an example, Figure S-2 (Supporting Information) shows a chromatogram extracted at 1072 cm^{-1} where two peaks corresponding to fructose and glucose can clearly be identified.

Results obtained using the PCA background correction approach are shown in Figure S-3 (Supporting Information). Figure S-3a (Supporting Information) shows the five PCA loadings calculated from the spectral information included in the **RSM** (i.e., calibration block). Again, the spectral features of the loadings agree well with the observed spectral changes during the LC gradient (see Figure S-1, Supporting Information). Then, using the score values obtained from the projection of the **SM** (317×372) spectra onto the space defined by the five PCA loading vectors, the background absorption is calculated as $\mathbf{TP}^T$, where **P** (372×5) is the k_{opt} loadings matrix and **T** (317×5) is the scores matrix of the **SM**. Background corrected data is shown in Figure S-3b (Supporting Information). The correlation coefficients between reference spectra and those extracted at the peak apex after background correction in the aforementioned range were 0.95 and 0.87 for fructose and glucose, respectively, being similar to those found using SIMPLISMA. From the plot of the calculated Q-residual values of the spectra included in both, the **RSM** and the **SM** matrix, the elution of the analytes can be easily identified (see Figure S-3c, Supporting Information).

MCR-ALS. As previously shown, in this case, both background correction methods provided similar results using five components. Results obtained using the SIMPLISMA approach have been chosen to show the performance of MCR-ALS analysis of background corrected data.

During MCR-ALS optimization, the model is fitted with a predefined number of components, k_{opt}, using initial estimates of either the **C** or the $\mathbf{S}^T$ matrix. Initially, log(*eigenvalues*) obtained

dx.doi.org/10.1021/ac2004407 |*Anal. Chem.* 2011, 83, 4855–4862

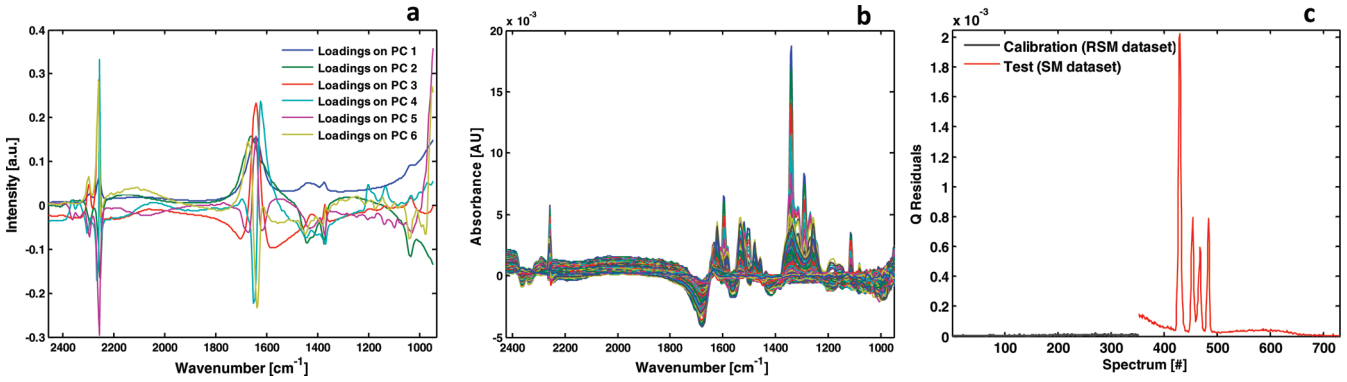


Figure 3. Results of the background correction of the data set 2 following the PCA approach. See text for details.

from a SVD analysis (see Figure 2a) and EFA analysis (results not shown) of the background corrected **SM** matrix indicated the possible presence of four components. The MCR-ALS initial spectral estimates of these four components were obtained using SIMPLISMA. Using these spectra, the corresponding concentration profiles of the four components in the corrected **SM** matrix were calculated by CLS (see Figure 2b). Whereas the concentration profiles calculated for components 1 and 3 showed LC peak shapes, this was not the case for components 2 and 4, which only seem to contain remaining background contribution.

MCR-ALS was performed using the following ALS constraints applied to components 1 and 3: non-negativity of the concentration profiles and unimodality. Besides, a convergence criterion of 0.02 and spectra normalization to equal area were employed. Unimodality forces the concentration profiles to have a single peak apex. Its application has no sense for resolution of residual solvent and instrumental drift contributions.[42] Under these conditions, convergence was achieved after 8 iterations, providing a percentage of explained variance of 90.2%. The time dependent concentration profiles obtained from MCR-ALS resolution are shown in Figure 2c. Concentration profiles of components 2 and 4 are noisy and showed anticorrelated trends related to residual solvent and background changes as well as to possible detector drifts. This was confirmed by spectral profiles shown in Figure 2d in which spectral features due to CO_2 and H_2O vapor were only present in spectra of components 2 and 4. Correlation coefficients between optimal MCR-ALS spectral profiles and reference spectra were 0.94 and 0.89 for fructose and glucose, respectively, in the 1396 and 1022 cm^{-1} range. Peak detection of components 1 and 3 could be improved, as compared to results found in a typical extracted chromatogram at a selected wavenumber (see Figure S-4, Supporting Information).

Data Set 2: Nitrophenol Analysis. *Background Correction.* In this second example, a standard mixture of four nitrophenols was analyzed by LC using a CH_3CN/H_2O solvent gradient. As in the previous example (see Figure 1), in Figure S-5a,b (Supporting Information), the log(*eigenvector*) values and the noise levels found as a function of the chemical rank are represented. From results shown, a k_{opt} value of 6 was selected for background correction. Corresponding spectra of the six components obtained by SIMPLISMA are represented in Figure S-5c (Supporting Information). Again, the extracted spectra are in agreement with spectral changes observed as a consequence of the LC gradient. The resulting background corrected data set is depicted in Figure S-5d (Supporting Information). A good compensation of the mobile phase contribution to the overall signal can be observed: a nonsloping baseline was obtained, and the four analytes were clearly identified. The chromatographic overlapping between components 2 and 3 is shown in Figure S-5e (Supporting Information) where a chromatogram extracted at 1338 cm^{-1} using a single point baseline correction at 1404 cm^{-1} is depicted. Remarkable low noise values were achieved throughout the whole spectrum. For example, noise values (measured as peak-to-peak) at 1800 cm^{-1} and 1640 cm^{-1} were <0.2 mAU and <3 mAU, respectively. Correlation coefficients between the spectra at the peak apex after background correction using the SIMPISMA approach and the reference spectra were 0.96, 0.93, 0.95, and 0.90 for 4-NP, 2-NP, 3m4-NP and 2,4-dN, respectively, in the spectral region from 1000 to 1700 cm^{-1}.

Results of the PCA background correction approach using 6 principal components to describe the **RSM** (352 × 392) data variance are summarized in Figure 3. Figure 3a shows the calculated PCA loadings (**P**), and in Figure 3b, the background corrected data matrix **SM** (379 × 392) is represented. Correlation coefficients between the spectra at the peak apex after background correction and the reference spectra were in this case 0.96, 0.92, 0.93, and 0.88 for 4-NP, 2-NP, 3m4-NP, and 2, 4-dN, respectively, using the same spectral range as before. In spite of the good background correction accuracy, a series of negative bands are present in the background corrected data. These spectral artifacts could not be eliminated using a different number of principal components to model the **RSM** data (data not shown). Again, the elution of the analytes can be accurately determined from the Q-residuals plot shown in Figure 3c.

MCR-ALS. MCR-ALS was applied to the background corrected data using the PCA approach. Here, six components were selected from the log(*eigenvalues*) plot obtained from SVD analysis of the background corrected SM matrix (see Figure 4a). Figure 4b shows the concentration profiles of the components obtained using SIMPLISMA and CLS which were subsequently used as initial estimates for the ALS resolution. A visual observation of the initial estimates helped us to conclude that components 1, 2, 4, and 5 correspond to the four analytes 4-NP, 2-NP, 3m4-NP, and 2,4-dN, respectively, and components 3 and 6 contained remaining spectral variation due to eluent gradient changes as well as changes in the instrumental conditions (e.g., detector drifts). Non-negativity and unimodality constraints (applied to components 1, 3, 4 and 5) were imposed to ALS optimization. Spectra normalization to equal area and a convergence criterion of 0.02 were also employed. Convergence was achieved after 8 iterations, with a lack of fit of 21% and a percentage of explained variance of 92%.

dx.doi.org/10.1021/ac2004407 |*Anal. Chem.* 2011, 83, 4855–4862

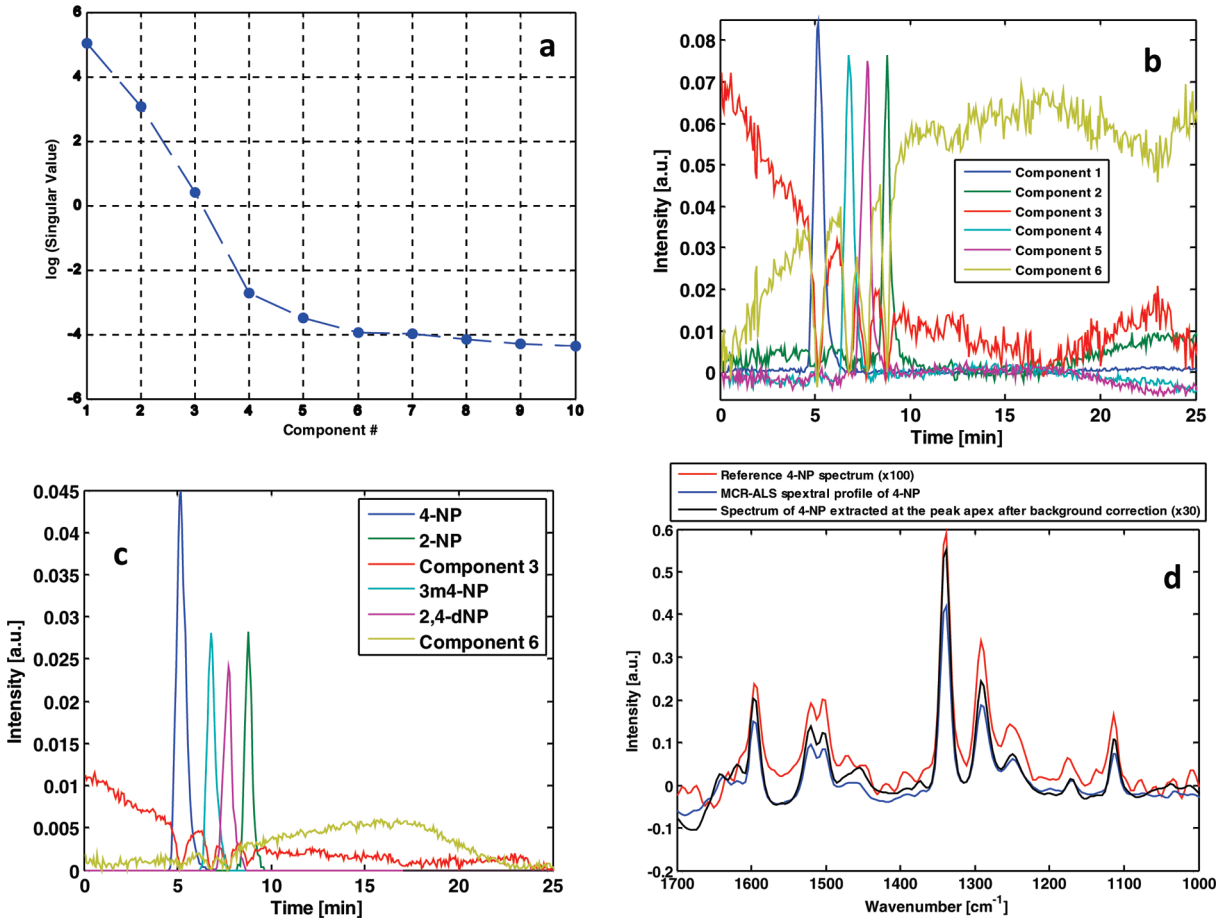


Figure 4. MCR-ALS analysis of the data set 2 after background correction: (a) selection of the chemical rank using SVD; (b) SIMPLISMA initial estimates of the concentration profiles; (c) calculated MCR-ALS concentration traces for the 6 components, and (d) calculated MCR-ALS, reference, and extracted spectra of 4-NP from a background corrected chromatogram.

Outcomes of the MCR-ALS optimization are shown in Figures 4c. Elution peaks of 4-NP, 2-NP, 3m4-NP, and 2,4-dN can be clearly seen in the MCR-ALS elution profiles shown in Figure 4c. Again, correlation coefficients between reference and MCR-ALS recovered spectra in the 1700 and 1000 cm^{-1} range were calculated to evaluate the spectral recoveries. Values found for 4-NP, 2-NP, 3m4-NP, and 2,4-dN after ALS resolution were 0.97, 0.92, 0.95, and 0.90, respectively. These correlation values were comparable to those obtained after MCR-ALS of the background corrected **SM** data matrix using the SIMPLISMA approach. In that case, ALS spectral recovery values obtained for 4-NP, 2-NP, 3m4-NP, and 2,4-dN after 3 iterations were 0.96, 0.93, 0.92, and 0.86, respectively, in the 1700 and 1000 cm^{-1} range.

As an example, the spectrum extracted from the peak apex of 4-NP of the background corrected data, its reference spectrum measured under stopped-flow conditions, and the MCR-ALS calculated spectrum are depicted in Figure 4d. Whereas after MCR-ALS analysis of background corrected data only minor changes in the spectral recovery values were found, ALS resolved concentration profiles showed, as previously, improved signal-to-noise ratios in comparison to those obtained from extracted chromatograms at selected wavenumbers.

It must be remarked that, like in the previous example, the use of further options of MCR-ALS (e.g., matrix augmentation, equality constraints for concentrations and/or spectra) was out of the scope of the present work. For quantitative and qualitative analysis, the use of multiple matrices of different mixtures at different concentration levels would likely improve the presented results. An evaluation of the use of these options and constraints for LC-IR data is still in progress and will be reported.

CONCLUSIONS

The contribution of background absorption to overall LC-IR signals could be compensated using two straightforward approaches based on PCA and SIMPLISMA and on the use of the spectral information obtained during the equilibration of the LC system. When MCR-ALS is applied directly to online gradient LC-IR data for the resolution of the spectra and concentration profiles of eluting analytes, good results are very difficult to obtain due to the high background contribution. On the contrary, when MCR-ALS was applied to background-corrected data, chemical rank selection and both peak and spectral resolution could be achieved. Besides, remaining signal variation due to background absorption and detector drift could also be compensated. Results found in online gradient LC-IR analysis of carbohydrates and nitrophenols as model examples illustrate well the potential of the assessed approach.

ASSOCIATED CONTENT

Supporting Information. Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

dx.doi.org/10.1021/ac2004407 |*Anal. Chem.* 2011, 83, 4855–4862

AUTHOR INFORMATION

Corresponding Author

*E-mail: guillermo.r.quintas@uv.es. Tel: +34 96 354 4838. Fax: +34 96 354 4845.

ACKNOWLEDGMENT

J.K. acknowledges the "V Segles" grant provided by the University of Valencia to carry out this study. Authors acknowledge the financial support of Ministerio de Educación y Ciencia (Projects AGL2007-64567 and CTQ2008-05719/BQU) and Conselleria d'Educació de la Generalitat Valenciana (Project PROMETEO 2010-055).

REFERENCES

(1) Armenta, S.; Lendl, B. *Anal. Bioanal. Chem.* **2010**, *397*, 297–308.
(2) Kuligowski, J.; Quintas, G.; de la Guardia, M.; Lendl, B. *Anal. Chim. Acta* **2010**, *679*, 31–42.
(3) Kuligowski, J.; Quintas, G.; Garrigues, S.; Lendl, B.; de la Guardia, M. *TrAC, Trends Anal. Chem.* **2010**, *29*, 544–552.
(4) http://www.daylightsolutions.com/; Daylight Solutions: San Diego, CA, USA, accessed May 2010.
(5) Brandstetter, M.; Genner, A.; Anic, K.; Lendl, B. *Analyst* **2010**, *135*, 3260–3265.
(6) Quintas, G.; Kuligowski, J.; Lendl, B. *Anal. Chem.* **2009**, *81*, 3746–3753.
(7) Griffiths, P. R.; de Haseth, J. A. *Fourier transform infrared Spectrometry*, 2nd ed.; John Wiley & Sons, Inc., San Francisco, USA, 2007.
(8) Kuligowski, J.; Carrion, D.; Quintas, G.; Garrigues, S.; de la Guardia, M. *J. Chromatogr., A* **2010**, *1217*, 6733–6741.
(9) Quintas, G.; Lendl, B.; Garrigues, S.; de la Guardia, M. *J. Chromatogr., A* **2008**, *1190*, 102–109.
(10) Wold, S.; Esbensen, K.; Geladi, P. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52.
(11) Maeder, M. *Anal. Chem.* **1987**, *59*, 527–530.
(12) Keller, H. R.; Massart, D. L. *Anal. Chim. Acta* **1992**, *263*, 21–28.
(13) Windig, W.; Guilment, J. *Anal. Chem.* **1991**, *63*, 1425–1432.
(14) Amrhein, M.; Srinivasan, B.; Bonvin, D.; Schumacher, M. M. *Chemom. Intell. Lab. Syst.* **1996**, *33*, 17.
(15) Harshman, R. A.; Lundy, M. E. *Comput. Stat. Data Anal.* **1994**, *18*, 39–72.
(16) Kiers, H. A. L.; Ten Berge, J. M. F.; Bro, R. *J. Chemom.* **1999**, *13*, 275–294.
(17) Bro, R.; Andersson, C. A.; Kiers, H. A. L. *J. Chemom.* **1999**, *13*, 295–309.
(18) Tauler, R. *Chemom. Intell. Lab. Syst.* **1995**, *30*, 133–146.
(19) de Juan, A.; Tauler, R. *J. Chromatogr., A* **2007**, *1158*, 184–195.
(20) Pere-Trepat, E.; Hildebrandt, A.; Barcelo, D.; Lacorte, S.; Tauler, R. *Chemom. Intell. Lab. Syst.* **2004**, *74*, 293–303.
(21) Pere-Trepat, E.; Lacorte, S.; Tauler, R. *Anal. Chim. Acta* **2007**, *595*, 228–237.
(22) Pere-Trepat, E.; Tauler, R. *J. Chromatogr., A* **2006**, *1131*, 85–96.
(23) Mas, S.; Fonrodona, G.; Tauler, R.; Barbosa, J. *Talanta* **2007**, *71*, 1455–1463.
(24) Edelmann, A.; Diewok, J.; Baena, J. R.; Lendl, B. *Anal. Bioanal. Chem.* **2003**, *376*, 92–97.
(25) Ruckebusch, C.; Vilmin, F.; Coste, N.; Huvenne, J. P. *Appl. Spectrosc.* **2008**, *62*, 791–797.
(26) Boelens, H. F. M.; Dijkstra, R. J.; Eilers, P. H. C.; Fitzpatrick, F.; Westerhuis, J. A. *J. Chromatogr., A* **2004**, *1057*, 21–30.
(27) Dijkstra, R. J.; Boelens, H. F. M.; Westerhuis, J. A.; Ariese, F.; Brinkman, U. A. T.; Gooijer, C. *Anal. Chim. Acta* **2004**, *519*, 129–136.
(28) Malinowski, E. R. *Anal. Chem.* **1977**, *49*, 612–617.
(29) Eilers, P. H. C. *Kwant. Meth.* **1988**, *23*, 25.
(30) Istvan, K.; Rajko, R.; Keresztury, G. *J. Chromatogr., A* **2006**, *1104*, 154–163.
(31) Bylund, D.; Danielsson, R.; Markides, K. E. *J. Chromatogr., A* **2001**, *915*, 43–52.
(32) Xu, C. J.; Liang, Y. Z.; Li, Y.; Du, Y. P. *Analyst* **2003**, *128*, 75–81.
(33) Meloun, M.; Capek, J.; Miksik, P.; Brereton, R. G. *Anal. Chim. Acta* **2000**, *423*, 51–68.
(34) Jackson, J. E.; Mudholkar, G. S. *Technometrics* **1979**, *21*, 9.
(35) Wise, B. M.; Gallagher, N. B.; Bro, R.; Shaver, J. M.; Windig, W.; Koch, R. S. *Chemometric tutorial for PLS_Toolbox and Solo*; Eigenvector Research, Inc.: Wenatchee, WA, 2006.
(36) Kuligowski, J.; Quintas, G.; Garrigues, S.; de la Guardia, M. *Talanta* **2008**, *77*, 779–785.
(37) Kuligowski, J.; Quintas, G.; Garrigues, S.; de la Guardia, M. *Talanta*, **2010**, *80*, 1771−1776.
(38) Jaumot, J.; Gargallo, R.; de Juan, A.; Tauler, R. *Chemom. Intell. Lab. Syst.* **2005**, *76*, 101–110.
(39) Gómez, V.; Miró, M.; Callao, M. P.; Cerdà, V. *Anal. Chem.* **2007**, *79*, 7767–7774.
(40) Kolhed, M.; Lendl, B.; Karlberg, B. *Analyst* **2003**, *128*, 2–6.
(41) Takamuku, T.; Tabata, M.; Yamaguchi, A.; Nishimoto, J.; Kumamoto, M.; Wakita, H.; Yamaguchi, T. *J. Phys. Chem. B* **1998**, *102*, 8880–8888.
(42) Pere-Trepat, E.; Lacorte, S.; Tauler, R. *J. Chromatogr., A* **2005**, *1096*, 111–122.