Available online at www.sciencedirect.com
SCIENCE DIRECT®
Analytica Chimica Acta 514 (2004) 203–209
ANALYTICA CHIMICA ACTA
www.elsevier.com/locate/aca

Parallel factor analysis of HPLC-DAD data for binary mixtures of lidocaine and prilocaine with different levels of chromatographic separation

Kent Wiberg [a,b], Sven P. Jacobsson [a,b,*]

[a] *AstraZeneca R&D Södertälje, Analytical Development, Pharmaceutical & Analytical R&D, SE-151 85 Södertälje, Sweden*
[b] *Department of Analytical Chemistry, University of Stockholm, SE-106 91 Stockholm, Sweden*

Received 9 October 2003; received in revised form 14 January 2004; accepted 23 March 2004

Abstract

A set of 17 samples containing a constant amount of lidocaine (667 μM) and a decreasing amount of prilocaine (667–0.3 μM) was analysed by LC-DAD at three different levels of separation, followed by parallel factor analysis (PARAFAC) of the data obtained. In Case 1 no column was connected, the chromatographic resolution (R_s) therefore being zero, while Cases 2 and 3 had partly separated peaks ($R_s = 0.7$ and 1.0). The results showed that in Case 1, analysed without any separation, the PARAFAC decomposition with a model consisting of two components gave a good estimate of the spectral and concentration profiles of the two compounds. In Cases 2 and 3, the use of PARAFAC models with two components resolved the underlying chromatographic, spectral and concentration profiles. The loadings related to the concentration profile of prilocaine were used for regression and prediction of the prilocaine content. The results showed that prediction of prilocaine content was possible with satisfactory prediction (RMSEP < 0.01). This study shows that PARAFAC is a powerful technique for resolving partly separated peaks into their pure chromatographic, spectral and concentration profiles, even with completely overlapping spectra and the absence or very low levels of separation.

Keywords: PARAFAC; HPLC-DAD; Binary mixtures; Lidocaine; Prilocaine; Partial separation

1. Introduction

High-performance liquid chromatography with diode array detection (HPLC-DAD) is a widely used analytical technique. DAD measures absorbance as a function of both time and wavelength, supplying a two-dimensional data matrix for every sample that is analysed. With DAD one spectrum is collected for every measuring point; in other words, one chromatogram is analysed at all wavelengths in contrast to the single wavelength detection HPLC, where only one chromatogram is recorded at a fixed wavelength. In recent years the demand for rapid HPLC analysis has increased in a number of analytical fields. If the speed of chromatographic analyses is increased, e.g. by using higher flows, shorter or smaller columns, etc., the chromatographic resolution generally becomes poorer and partially separated peaks often occur. However, by analysing the DAD data of unresolved peaks with chemometric tools, it is possible to resolve them. Examples of these methods are evolving factor analysis (EFA) [1,2], window factor analysis (WFA) [3], heuristic evolving latent projections (HELP) [4,5], eigenstructure tracking analysis (ETA) [6] and multivariate curve resolution [7]. All these methods use principal component analysis (PCA) for the decomposition of the two-way LC-DAD data. If a number of samples are analysed, the resulting data are three-dimensional matrices consisting of absorbance as a function of time, wavelength and samples. These three-way data sets can be unfolded, whereupon bilinear chemometric methods like PCA can be applied. However, if the three-way matrix is to be analysed without unfolding, other methods must be used. Examples of chemometric methods for decomposition of multiway data are parallel factor analysis (PARAFAC) [8] and *n*-way partial least squares [9].

* Corresponding author. Tel.: +46-8-553-289-68; fax: +46-8-553-277-30.
E-mail address: sven.jacobsson@astrazeneca.com (S.P. Jacobsson).

0003-2670/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2004.03.062

A drawback of PCA, however, is that on matrices with a chemical rank larger than one it is not possible to obtain the pure chromatographic and spectral profiles, although by analysing three-way chromatographic data for partially separated peaks with PARAFAC, this becomes possible [10–12]. The aim of the present study was to investigate the ability of PARAFAC to resolve the pure spectral, chromatographic and concentration profiles of some partially separated peaks at different levels of separation, something which to our knowledge has not previously been reported in the literature. PARAFAC was used for the analysis of HPLC-DAD data for binary mixtures with an absence or a low level of chromatographic resolution. Three levels of separation were used, ranging from a chromatographic resolution (R_s) of 0–1.0. The local anaesthetic drugs lidocaine and prilocaine were used as model substances. These two compounds absorb in the same wavelength range of 245–290 nm and thus show completely overlapping UV spectra.

2. Theory

PARAFAC is a decomposition method for three-way arrays (or higher) that can be seen as a generalisation of bilinear PCA to higher order arrays [8]. In PARAFAC each component is trilinear, in contrast to bilinear PCA, where one score and one loading vector are obtained for each component. Three loading vectors (a_{if}, b_{jf}, c_{kf}) are therefore given for each PARAFAC component (F). The PARAFAC model of a three-way array is found by minimising the sum of the squares of the residuals e_{ijk}:

$$x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk} \quad (1)$$

The decomposition of a three-way array of HPLC-DAD data is schematically described in Fig. 1. Each PARFAC component gives three loadings: one related to the chromatographic profile (a_f), one related to the spectral profile (b_f) and one related to the content of the samples (c_f). Hence one loading is given for each dimension in the data. Alternating least squares (ALS) can be used to find the solution to the PARAFAC model. A special feature of PARAFAC compared to PCA is that the solutions obtained are unique, which means that there is no rotational problem as in PCA. Hence the loadings obtained with PARAFAC can be directly interpreted chemically are directly chemically. Furthermore, the PARAFAC algorithm is not sequential as in PCA since the trilinear model is calculated simultaneously for all components. Different methods of determining the correct number of components have been suggested. Examples are split-half [8], core consistency diagnostic [13,14] or the use of external knowledge of the data being modelled. In this study the latter approach was used, through examination of the PARAFAC loadings obtained.

3. Experimental

3.1. Instrumentation

The HPLC system used was a Dionex Summit HPLC system with Chromeleon, Version 6.11 as chromatographic data system. This HPLC system consisted of a PDA 100 diode array detector, an ASI 100T autosampler and a P580 HPG pump (flow 1.0 ml min^{-1}). The chromatographic column used was a μBondapak C18 and the mobile phase was phosphate buffer with a varying amount of acetonitrile. The samples were prepared by weighing into stock solutions with a calibrated balance, Sartorius MC5. An automatic diluter, Hamilton Microlab 1000, was then used to obtain the different concentrations used in the study. All the calculations were made with Matlab, Version 6.0 with PLS toolbox 2.1.

3.2. Reagents

Throughout the study spectroscopic-quality chemicals were used. The phosphate buffer used was obtained by mixing 1.3 ml of 1 M sodium dihydrogen phosphate with 32.5 ml of 0.5 M disodium hydrogen phosphate with water. Analytical grade acetonitrile was used. AstraZeneca Bulk Production, Södertälje, supplied the lidocaine hydrochloride and prilocaine hydrochloride. A Millipore Milli-Q filtration/purification system was used to produce the water used in the study. The structures and UV spectra of lidocaine and prilocaine are shown in Fig. 2, where it can be seen that the absorbance of the two compounds is in the same wavelength region of 245–290 nm.

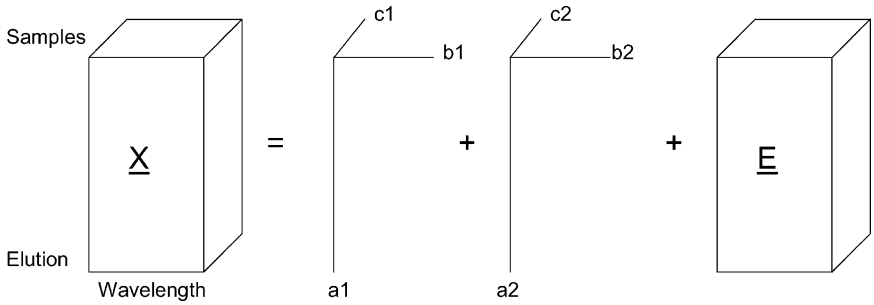

Fig. 1. Schematic description of the decomposition of a three-way matrix ***X*** with PARAFAC. Two components are calculated.

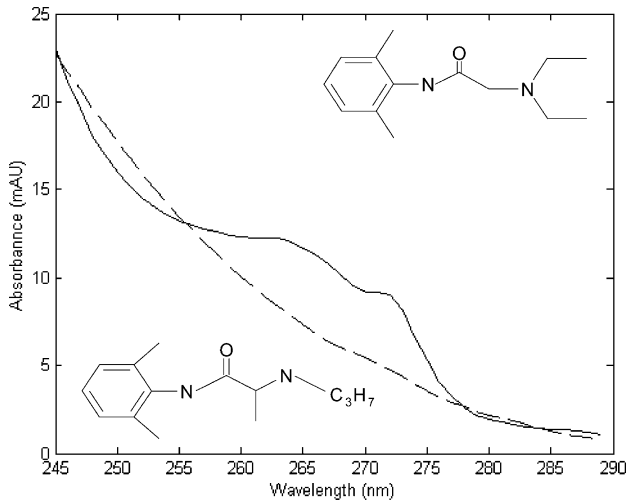

Fig. 2. Structural formula and UV spectra (245–290 nm) of lidocaine (667 μM) and prilocaine (130 μM). The solid line and the upper right structural formula relate to lidocaine and the dashed line and lower left structure to prilocaine.

Table 1
Prilocaine concentration in the samples used in the study

Sample no.	Prilocaine (μM)	Amount of prilocaine (conc. %)
X1	667	50.0
X2	533	44.4
X3	400	37.5
X4	267	28.6
X5	133	16.7
X6	66.7	9.1
X7	40.0	5.7
X8	26.7	3.9
X9	13.3	2.0
X10	6.7	1.0
X11	5.3	0.8
X12	4.0	0.6
X13	2.7	0.4
X14	1.3	0.2
X15	0.7	0.1
X16	0.4	0.06
X17	0.3	0.04

These 17 binary mixtures contained 667 μM lidocaine and varying amounts of prilocaine.

3.3. Analytical procedure

DAD data from three cases of chromatographic separation were used in this study: Case 1, $R_s = 0$ (no separation at all); Case 2, $R_s = 0.7$; Case 3, $R_s = 1.0$ (partly separated peaks). In Case 1 the analysis was carried out without any chromatographic column connected, instead a back-pressure tube (PEEK, i.d. 0.1 mm, length 6 m) being connected between the autosampler and the detector to provide pressure for the pump to work against. In Cases 2 and 3 an ordinary chromatographic column was employed and a varying amount of acetonitrile in the mobile phase gave the different separations. At all three levels of separation a test set of 17 samples (X1–X17) containing binary mixtures of lidocaine and prilocaine was analysed. These 17 samples contained 667 μM lidocaine and a decreasing amount of prilocaine (Table 1). A three-dimensional plot of the absorbance as a function of retention time and wavelength in the DAD data for sample X5, Case 2 $R_s = 0.7$, is shown in Fig. 3.

When the concentrations of the two compounds are roughly equal, the UV spectra of lidocaine and prilocaine are clearly visible, as can be seen in Fig. 3.

The wavelength region used was 245–290 nm, with a resolution of 1 nm and a data collection rate of 10 Hz. The runtimes used were 0.5 min for Case 1 and 7 min for Cases 2 and 3. These analyses gave data matrices of $4200 \times 45 \times 17$ for Cases 2 and 3 and $300 \times 45 \times 17$ for Case 1. These three-way data matrices were then decomposed with PARAFAC. Different PARAFAC models were calculated with between one and four components and the loadings and residuals obtained were used to find the appropriate number of components to use in each case. No scaling or centring took place before modelling. The wavelength region chosen has previously been used in a number of studies [15–17]. The runtime chosen in Case 1 was only 0.5 min since the sample reached the detector cell after about 0.18 min. In Case 1, where no chromatographic column was used, some small shifts in the retention time of the single peak in the 17 samples were observed. Since the PARAFAC1 algorithm used in this study is sensitive to retention time shifts [14], corrections were made to the shifts of the samples in Case 1.

In Fig. 4 the chromatograms for samples X3, X5 and X10 in Cases 1–3 with detection at 245 nm are shown. For samples X1–X4 in Cases 2 and 3, the height of the two partly separated peaks is higher than about 30 mAU since the summed absorbance of lidocaine and prilocaine at 245 nm is high. Prilocaine also has a higher extinction coefficient than lidocaine. However, when the prilocaine content of the samples decreases in samples X5–X17, the height of the lidocaine peak is constant, while the prilocaine peak decreases in intensity. In sample X10 it becomes hard to see the small prilocaine peak in front of the lidocaine peak in the chromatogram (Fig. 4b and c), and from sample X12 onward it is no longer possible to see the small prilocaine peak.

4. Results and discussion

In Fig. 5 the loadings from the PARAFAC decomposition of the data for samples X1–X17, analysed without any chromatographic separation (Case 1, $R_s = 0$), are shown. The model used contained two PARAFAC components that explained >99.98% of the variation in the data. The loadings related to the chromatographic profile are shown in Fig. 5a (a1, a2), the spectral profile in Fig. 5b (b1, b2) and the concentration profile in Fig. 5c (c1, c2). This notation of the loadings refers to the notation used in Eq. (1) and in Fig. 3. When no chromatographic column is connected, the resulting ‘chromatogram’ contains only a single peak since no separation is applied. In Fig. 5a the loadings a1 (dashed line)

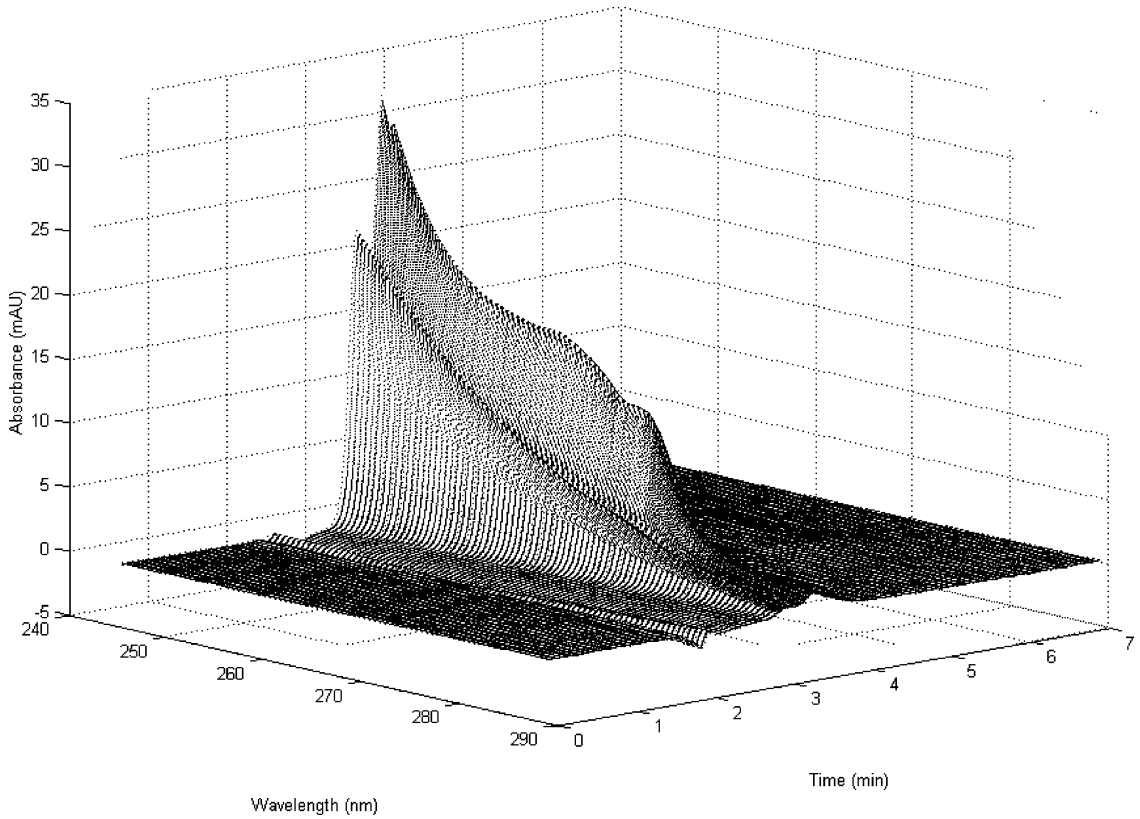

Fig. 3. Three-dimensional plot of the DAD data for sample X5, Case 2, $R_s = 0.7$. The absorbance is shown as a function of wavelength and retention time.

and a2 (solid line) are shown and, as can be seen, there is a complete overlap of the chromatographic profiles of the two compounds. In Fig. 5b the spectral profiles are shown and, as can be seen, the PARAFAC decomposition has given a good estimate of the UV spectrum of lidocaine (solid line) and of prilocaine (dashed line). Moreover, the concentration profiles of the two compounds have been found as shown in Fig. 5c, where the solid line related to lidocaine shows the constant lidocaine amount in all samples and the dashed line the decreasing prilocaine concentration. The two PARAFAC components were thus able to give a good estimate of the spectral and concentration profiles despite the fact that no separation was applied in the analysis. PARAFAC decomposition of the data for the samples in Case 1 was also carried out without any correction to the small shifts in retention time in some samples, although these results were biased and more PARAFAC components were needed to give an estimate of the underlying profiles (results not shown here). From these results it can therefore be concluded that if significant shifts in the retention time occur, corrections to them should be made or PARAFAC2 [13,14] should be used for the analysis.

The loadings from the PARAFAC decomposition of the data for samples X1–X17 with partly separated peaks (Case 2, $R_s = 0.7$) are shown in Fig. 6. Two PARAFAC components have been calculated and 99.98% of the variation in the data has been explained. The loadings related to the chromatographic profile are shown in Fig. 6a (a1, a2), the spectral profiles in Fig. 6b (b1, b2) and the concentration profile in Fig. 6c (c1, c2). As can be seen in Fig. 6a, the prilocaine peak elutes before the lidocaine peak, while in Fig. 6b it can be clearly seen that the second loadings (b1 and b2) have captured the true spectral profiles of the two compounds. Comparing the spectral profile in Fig. 6b with the corresponding profile found in Case 1 (Fig. 5b), it can be seen that when a small level of separation was applied (Case 2), a more correct representation of the true UV spectra of the two compounds was found. In Fig. 5b the spectral profiles of lidocaine and prilocaine appear to be somewhat more mixed and less distinct, which can probably be explained by the absence of separation of the two compounds in Case 1. In Fig. 6c the loadings c1 and c2 show the concentration profiles of lidocaine (solid line) and prilocaine (dashed line) for samples X1–X17 in Case 2. As can be seen, the PARAFAC loadings have found the constant lidocaine content in all samples as well as the decreasing prilocaine content. Comparing this figure with the corresponding figure in Case 1 (Fig. 5c), it can be seen that the level of the lidocaine content (solid line) differs in the two figures. The explanation for this also probably lies in the absence of separation in Case 1. Since the raw data in Case 1 consist of a single peak with only a varying peak height, the constant lidocaine concentration obtains higher loadings compared to when a small level of separation is applied in Case 2.

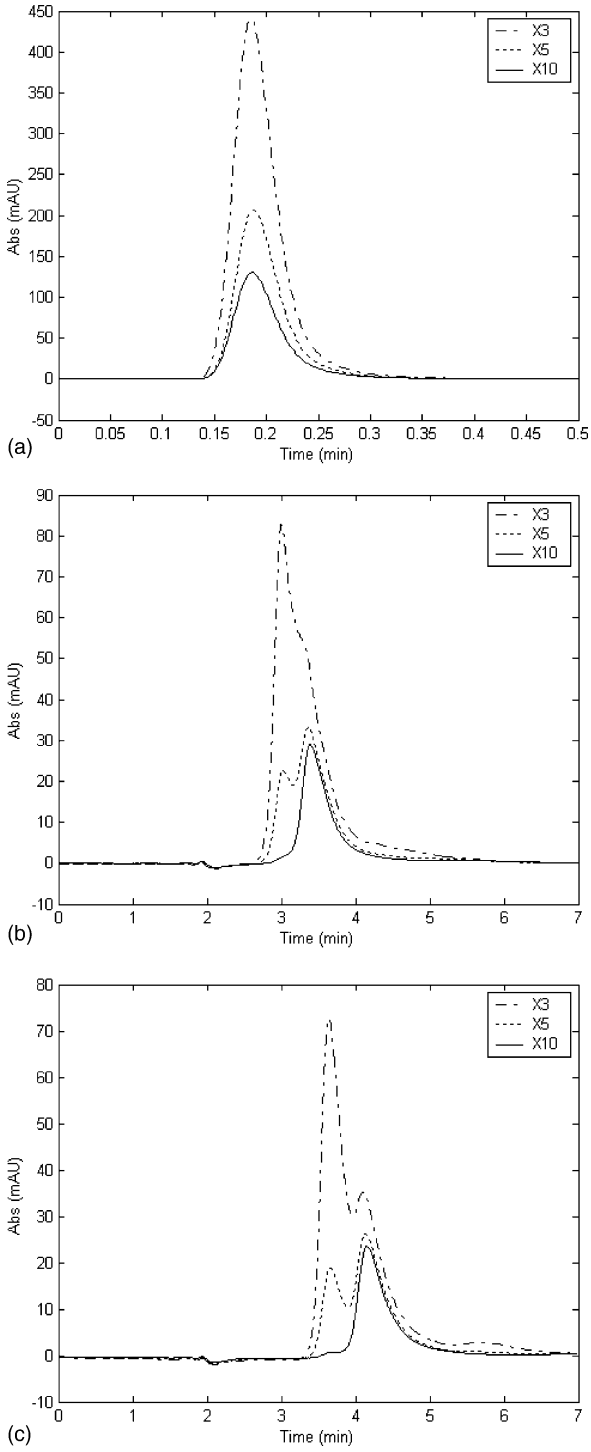

Fig. 4. Chromatograms at 245 nm for samples X3, X5 and X10, Cases 1–3: (a) Case 1, $R_s = 0$, (b) Case 2, $R_s = 0.7$, (c) Case 3, $R_s = 1.0$.

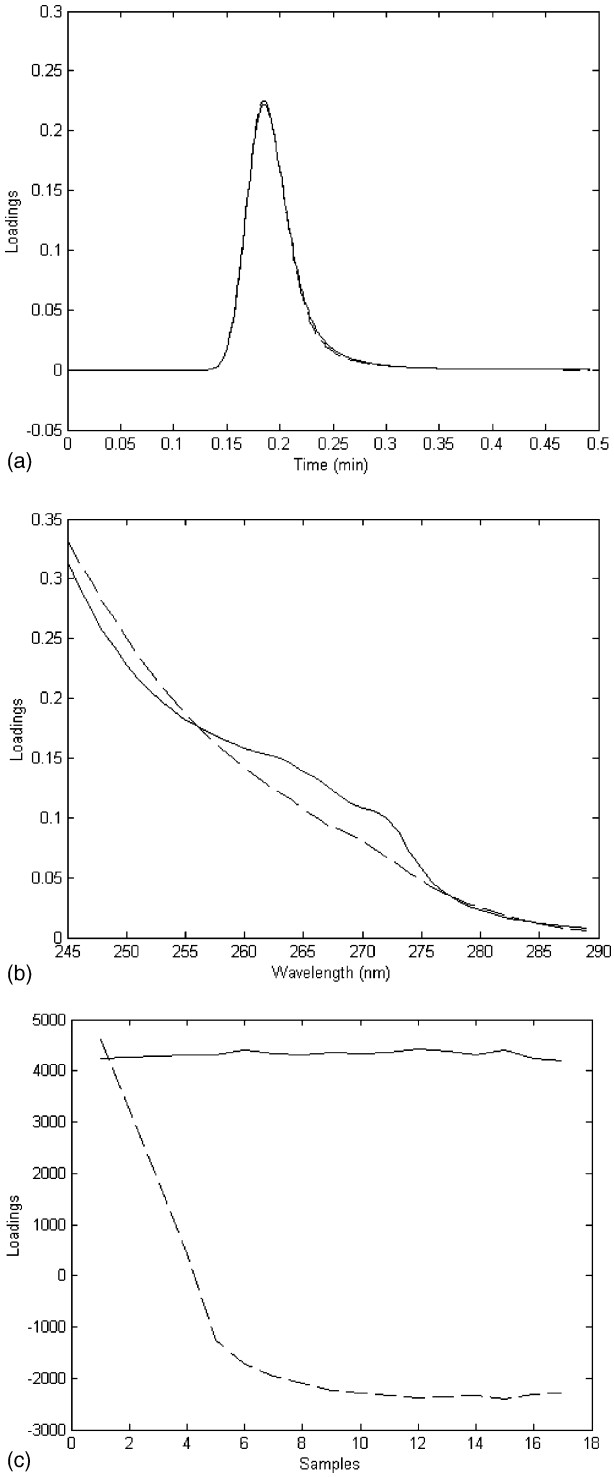

Fig. 5. Loadings from PARAFAC decomposition with a two-component model of samples X1–X17, Case 1 ($R_s = 0$), loadings from the first loadings from the first PARAFAC component being shown by a dashed line and loadings from the second component by a solid line: (a) loadings a1 and a2; (b) loadings b1 and b2; (c) loadings c1 and c2.

In Fig. 7 the analogous loadings of the PARAFAC decomposition of the data for samples X1–X17 with partly separated peaks Case 3 ($R_s = 1.0$) are shown. A PARFAC model with two components was used and explained 99.92% of the variation in the data. As can be seen, the results are the same as in Fig. 6, the chromatographic, the spectral and the concentration profile clearly being given by the PARAFAC loadings.

PARAFAC models were also tested for Cases 1–3 with different numbers of components, although none of

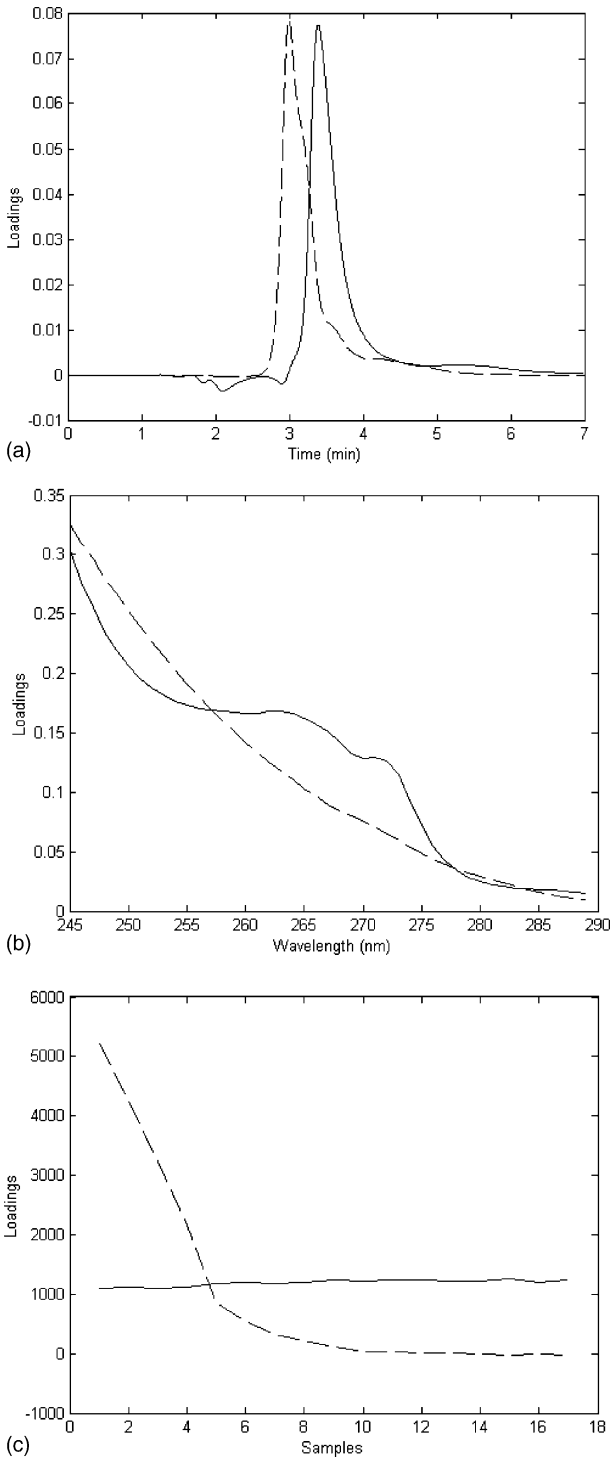

Fig. 6. Loadings from PARAFAC decomposition with a two-component model on samples X1–X17, Case 2 ($R_s = 0.7$), loadings from the first PARAFAC component being shown by a dashed line and loadings from the second component by a solid line: (a) loadings a1 and a2; (b) loadings b1 and b2; (c) loadings c1 and c2.

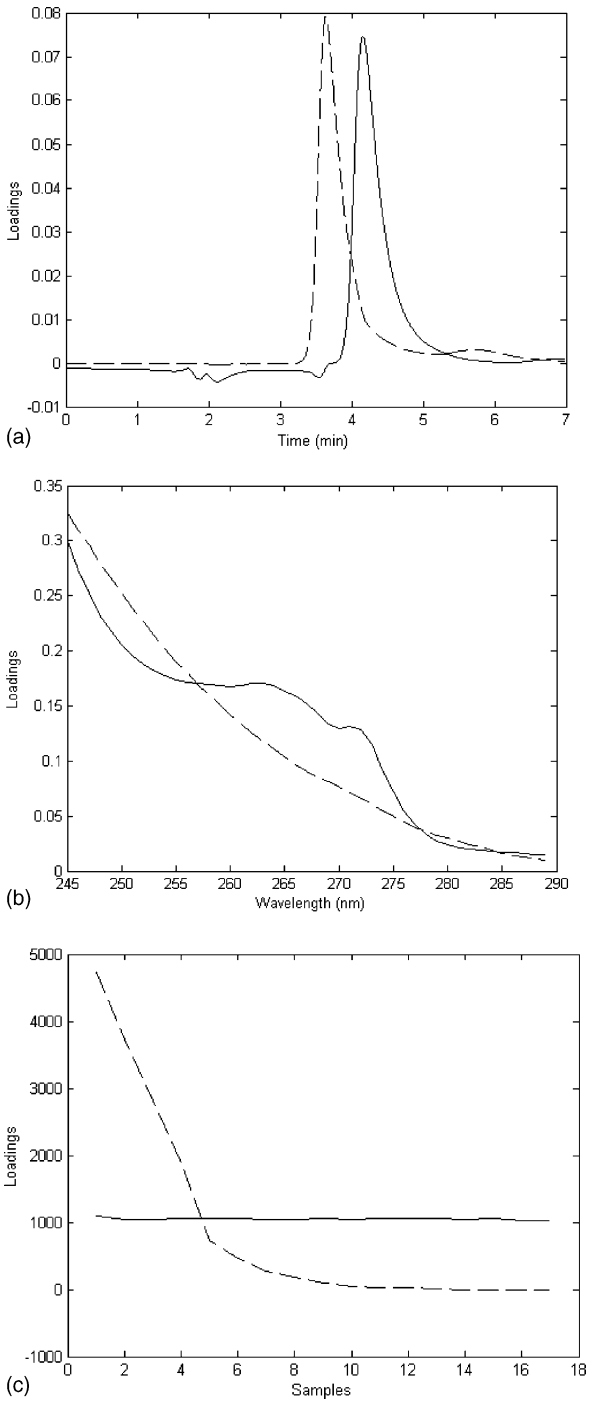

Fig. 7. Loadings from PARAFAC decomposition with a two-component model of samples X1–X17 Case 3 ($R_s = 1.0$), loadings from the first PARAFAC component being shown by a dashed line and loadings from the second component by a solid line: (a) loadings a1 and a2, (b) loadings b1 and b2, (c) loadings c1 and c2.

these models improved the interpretation of the loadings. PARAFAC models were also made from samples X5–X17 only, where the height of the lidocaine peak is constant and the prilocaine peak is smaller than the lidocaine peak for all samples. However, the results obtained were the same as with the PARAFAC models using all 17 samples.

To test the ability to use the result of the PARAFAC decomposition for determination of prilocaine, the loadings describing the prilocaine content (c1) were regressed against the true prilocaine concentration. Samples X1–X17 were di-

Table 2
Prediction results (RMSEP values in μM) for the determination of prilocaine in sample sets I and II

	Case 1, $R_s = 0$	Case 2, $R_s = 0.7$	Case 3, $R_s = 1.0$
(a) Calibration with the samples in sample set I and prediction of the samples in sample set II	4.3	5.1	3.4
(b) Calibration with the samples in sample set II and prediction of the samples in sample set I	9.5	9.9	9.5

vided into two sample sets, I and II. Sample set I contained the odd numbered samples X1, X3, ..., X17, while sample set II consisted of the even numbered samples X2, X4, ..., X16. The two sample sets were then used both as a calibration set and a test set. Firstly, sample set I was used for calibration and sample set II as the test set. This was done for Cases 1–3 and the prediction results obtained (RMSEP) are shown in Table 2a. Secondly, the samples in sample set II were used for calibration and sample set I for testing the predictions. This was also done for all three cases of different separation and the results are shown in Table 2b. As can be seen in this table, the prediction results obtained were good, with RMSEP values $<10\,\mu M$. It was therefore possible to use the loadings obtained with PARAFAC decomposition for the determination of prilocaine. This was valid for the data obtained without any chromatographic separation as well as with a low level of separation.

Factors that might affect the results of using PARAFAC decomposition on HPLC-DAD data by causing departures from tri-linearity are shifts in retention time and changes of peak shape. When retention time shifts are present the PARAFAC2 algorithm can be applied [14] and if also changes in peak shape are present, multivariate curve resolution-alternating least squares (MCR-ALS) [18] might be an alternative.

5. Conclusion

This study shows that PARAFAC is a powerful chemometric technique for resolving partly separated peaks into the pure chromatographic, spectral and concentration profiles even with completely overlapping spectra and the absence or very low levels of separation. When a small level of separation was applied, the PARAFAC decomposition found the true underlying chromatographic, spectral and concentration profiles of the binary samples. For samples analysed without any chromatographic separation, the PARAFAC decomposition gave a good estimate of the spectral and concentration profiles. It was also possible to use the loadings from the PARAFAC decomposition for regression and determination of prilocaine. It should therefore be possible to implement PARAFAC in impurity profiling, leading to more rapid analyses and higher throughput.

References

[1] M. Maeder, A.D. Zuberbuehler, Anal. Chim. Acta 181 (1986) 287.
[2] M. Maeder, Anal. Chem. 59 (1987) 527.
[3] E.R. Malinowski, J. Chemomet. 6 (1992) 29.
[4] O.M. Kvalheim, Y. Liang, Anal. Chem. 64 (1992) 936.
[5] Y. Liang, O.M. Kvalheim, H.R. Keller, D. Massart, P. Kiechle, F. Erni, Anal. Chem. 64 (1992) 946.
[6] J. Toft, O.M. Kvalheim, Chemomet. Intell. Lab. Syst. 19 (1993) 65.
[7] R. Tauler, D. Barceló, Trends Anal. Chem. 12 (1993) 319.
[8] R. Bro, Chemomet. Intell. Lab. Syst. 38 (1997) 149.
[9] R. Bro, J. Chemomet. 10 (1996) 47.
[10] J.L. Beltrén, J. Guiteras, R. Ferrer, J. Chromatogr. A 802 (1998) 263.
[11] P. Hindmarch, K. Kavianpour, R.G. Brereton, Analyst 122 (1997) 871.
[12] H.L. Wu, M. Shibukawa, K. Oguma, J. Chemomet. 12 (1998) 1.
[13] H.A.L. Kiers, J.M.T. Ten Berge, R. Bro, J. Chemomet. 13 (1999) 275.
[14] R. Bro, C.A. Andersson, H.A.L. Kiers, J. Chemomet. 13 (1999) 295.
[15] K. Wiberg, A. Hagman, P. Burén, S.P. Jacobsson, Analyst 146 (2001) 1142.
[16] K. Wiberg, A. Hagman, S.P. Jacobsson, J. Pharmaceut. Biomed. Anal. 30 (2003) 1575.
[17] K. Wiberg, M. Andersson, A. Hagman, S.P. Jacobsson, J. Pharmaceut. Biomed. Anal. 33 (2003) 859.
[18] A. de Juan, R. Tauler, J. Chemomet. 15 (2001) 749.